# FINAL REPORT

# Bandwidth Reduction of Sleep Information

## Vol. II.

By

A. J. Welch, Philip C. Richardson, Jane Mockford

and Joanne M. Aldredge

Department of Electrical Engineering

Technical Report No. 115

November 1, 1971

BIO-MEDICAL ENGINEERING RESEARCH LABORATORY

prepared under Contract F41609-71-C-0002

Brooks Air Force Base, San Antonio, Texas



ELECTRONICS RESEARCH CENTER

THE UNIVERSITY OF TEXAS AT AUSTIN

Austin, Texas 78712

Unclassified
Security Classification

# DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

| | |
|---|---|
| 1. ORIGINATING ACTIVITY (Corporate author)<br>Bio-Medical Engineering Research Laboratory<br>Electronics Research Center<br>The University of Texas at Austin, Austin, Texas | 2a. REPORT SECURITY CLASSIFICATION<br>Unclassified<br>2b. GROUP |

3. REPORT TITLE

BANDWIDTH REDUCTION OF SLEEP INFORMATION. VOL. II.

4. DESCRIPTIVE NOTES (Type of report and inclusive dates)

Final Report Sept 70 - Nov 71

5. AUTHOR(S) (First name, middle initial, last name)


A. J. Welch
Philip C. Richardson
Jane Mockford
Joanne M. Aldredge


| 6. REPORT DATE | 7a. TOTAL NO. OF PAGES | 7b. NO. OF REFS |
|---|---|---|
| November 1971 | 88 | 25 |

| | |
|---|---|
| 8a. CONTRACT OR GRANT NO.<br>F41609-71-C-0002<br>b. PROJECT NO.<br>c. Task No. 631903<br>d. | 9a. ORIGINATOR'S REPORT NUMBER(S)<br>Technical Report No. 115<br>9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) |

10. DISTRIBUTION STATEMENT

This document has been approved for public release and sale; its distribution is unlimited.

| 11. SUPPLEMENTARY NOTES | 12. SPONSORING MILITARY ACTIVITY |
|---|---|
| | USAF School of Aerospace Medicine<br>Aerospace Medical Division (AFSC)<br>Brooks Air Force Base, Texas |

13. ABSTRACT


This report discusses the possibility of extracting sleep information from heart rate data. Eight hours of sleep EEG, EOG and electrocardiograms were recorded on FM magnetic tape for two nights for each of eight subjects. The time in milliseconds between heart beats was written on digital magnetic tape. The data were grouped into records containing 128 consecutive beat-to-beat intervals and a large number of descriptors was computed for each record. These descriptors for each record were the mean value, the sample variance, the nine-interval histogram of Z scores, and a set of Fourier transforms. Analysis of variance was used as a general guide to descriptor significance for each subject. The discriminant analysis procedure described by Rao and popularized by Cooley and Lohnes was used to sleep stage classify heart rate data. Accuracy of the procedure was determined in terms of percent correct classifications, correlation coefficient of the computerized sleep pattern with respect to the EEG hand scored pattern, and an empirically derived cost function.

The results of this study suggest that for a single night of sleep a reasonable accuracy of sleep stage classification is possible. However the variability in heart rate from night-to-night for any one individual produces unacceptably poor classification results on the second night.


DD FORM 1 NOV 65 1473

Unclassified
Security Classification

Unclassified
Security Classification

| 14. KEY WORDS | LINK A | | LINK B | | LINK C | |
|---|---|---|---|---|---|---|
| | ROLE | WT | ROLE | WT | ROLE | WT |
| | | | | | | |

Unclassified
Security Classification

# BANDWIDTH REDUCTION OF SLEEP INFORMATION

VOL. II.

For Vol. I
See AD-718125 (Pg. 2i)

By

A.J. Welch, Philip C. Richardson, Jane Mockford

and Joanne M. Aldredge

Final Report for

Brooks Air Force Base

Contract F41609-71-C-0002

November 1, 1971

BIO-MEDICAL ENGINEERING RESEARCH LABORATORY

ELECTRONICS RESEARCH CENTER

The University of Texas at Austin

Austin, Texas 78712

ABSTRACT

This report discusses the possibility of extracting sleep information from heart rate data. The recognition of sleep stages or even the ability to differentiate sleep from wakefulness using heart rate information alone rather than the conventional EEG measures could expend the scope of sleep studies. In many situations where it is desirable to evaluate wakefulness and sleep, EEG electrodes become unreliable after a few days and the time bandwidth requirements of recording and transmitting the EEG are excessive.

Eight hours of sleep EEG, EOG and electrocardiograms were recorded on FM magnetic tape for two nights. The method of data collection and sleep scoring of the EEG was reported in detail by Lessard, Ford and Hughes of the USAF School of Aerospace Medicine (14). Copies of the FM magnetic tapes containing sleep data for ten subjects were supplied to The University of Texas Bio-Medical Engineering Laboratory by the U. S. Air Force School of Aerospace Medicine.

Double differentiation of the filtered electrocardiogram and threshold logic units were used to detect the peak of each R-wave. The time in milliseconds between heart beats was written on digital magnetic tape. The data were grouped into records containing 128 consecutive beat-to-beat intervals and eleven descriptors were computed for each record. These descriptors for each record were the mean value $\overline{X}$, the sample variance $S^2$, and the nine-interval

histogram of the beat-to-beat R-R intervals. We represented each R-R interval as $X_i$ and the mean value and standard deviation of the 128 R-R intervals as $\bar{X}$ and S, respectively; then the standardized value for the ith heart beat was

$$Z_i = \frac{X_i - \bar{X}}{S} \qquad (1)$$

The distribution of Z scores had a mean value of zero and standard deviation of one. Histogram intervals were one standard deviation wide and the number of standardized scores falling in each of the nine intervals were used as the value of the histogram descriptors. Each group of 128 heart intervals had a different mean value, $\bar{X}$ and sample standard deviation, S.

The outer intervals of the histogram extended from $-\infty$ to $-1\text{-}3/4$ and $+1\text{-}3/4$ to $+\infty$.

Analysis of variance was used to determine descriptor significance for each subject. This procedure tested the null hypothesis for each descriptor. In other words, what was the probability that a descriptor mean value for awake and the five stages of sleep were equal? Our program used an F-test to compute the probability that the mean values of the descriptors for each stage of sleep were equal. The hypothesis was rejected if $P < .01$.

The discriminant analysis procedure described by Rao (17) and popularized by Cooley and Lohnes (7) was used to sleep stage classify heart rate data. Approximately one-half of the first recorded night of sleep for each individual was used as a training set in the discriminant analysis procedures. Once the training set had been obtained both nights of sleep for the individual were sleep stage classified into awake and stage 1, 2, 3, 4, and REM sleep. Accuracy of the procedure was determined in terms of percent correct classifications, correlation coefficient of the computerized sleep pattern with respect to the EEG hand scored pattern and an empirically derived cost function.

The results of this study suggest that for a single night of sleep a reasonable accuracy of sleep stage classification is possible. However the variability in heart rate from night-to-night for any one individual produces unacceptably poor classification results on the second night.

## TABLE OF CONTENTS

## FIGURES

TABLES

TABLES (Continued)

TABLES (Continued)

## INTRODUCTION

Extended periods of sleep deprivation commonly produces a decrease of performance capabilities at skilled tasks (4,10,13,18) in addition to unhealthy changes in personality profiles (22,23). Indeed, hallucinations have been observed in laboratory experiments (9). Kleitman (12) reports that "among the effects of prolonged wakefulness are irritability and mental disorganization, leading to daydreaming and automatic behavior, occasionally bordering on temporary insanity." Decreased performance prevents a person from meeting the requirements of many military situations which require maximal alertness and performance by the on-duty personnel. Berry (4) reports that fatigue due to inadequate rest interferred with the ability of the astronauts to perform tasks in the Apollo VII and VIII missions. Usually the state of alertness and performance for an individual is associated with the amount of rest and sleep he has obtained.

Unfortunately the technical difficulties in obtaining sleep information have impeded sleep research outside the laboratory. Classically sleep is evaluated from electroencephalographic data (EEG). The instability of the EEG electrodes over extended periods of time and the lack of an automated process for evaluation of the EEG has discouraged meaningful research of sleep in military situations.

One possible solution to this problem would be the development of an alternate source of sleep information. The source we consider in this report is beat-to-beat heart rate. This electrophysiological measurement is more stable than the EEG over long periods of time and there is evidence that average heart rate is influenced by sleep. It is the intent of this report to determine the feasibility of using instantaneous heart rate in an automated process as an indicator of the sleep-wakefulness cycle.

## Statement of Problem

Pilots in flight are not always able to report accurately their physical condition, particularly with respect to drowsiness-wakefulness, which affects alertness and operational capability. A simple objective measure of this condition under operational conditions is desirable. Any solution must keep the sensory system simple and must keep the required transmission bandwidth small. One possibility is to derive this wakefulness information from beat-to-beat heart rate data, transmit the heart rate data to ground stations, and use computer analysis to determine sleep stage from heart rate derived measures.

In a previous report from The University of Texas at Austin (21) we described the computation of several different measures of beat-to-beat heart rate. We also computed the possible utility of each of these measurements to a sleep stage classification program. This report examines computer classification of sleep stages utilizing the measures described in our previous report.

Review of Literature

Since the classical work of Dement and Kleitman (8), the depth and duration of sleep has been determined by examination of electroencephalographic data. However, the recording of a full night of sleep EEG on a strip chart recorder results in a bulky set of data. Further the visual interpretation of these data is a time-consuming task since an individual must manually scan up to 1,000 feet of strip chart record! Trained personnel, using well-documented criteria to score the sleep records, have not produced a consistent procedure for scoring a night of sleep with a guaranteed accuracy better than 90% (24). Monroe (16) reports inter-rater consistency in scoring sleep records by different specialists to be only 65 percent. Many inaccuracies may be due to the marked degree of subjectivity that must be used in visually scanning lengthy records. In some situations such as space flight, bandwidth, weight considerations and poor electrode techniques suggest that an alternate signal to the EEG is needed.

Coupling between sleep activity in the brain and autonomic nervous system motor activity has been documented by many investigators (5,6,11,15,19).

Snyder (20) reported significant changes ($P < .05$) in average levels of blood pressures, respiration and heart rate between stages 1-REM and stage 2. Data were recorded from twelve subjects for a total of thirty nights. There was a 6% average increase in heart rate, a 7% average increase in respiratory rate and a 4% average increase in systolic blood pressure from stage 2 to stage 1-REM. Significant changes ($P < .05$) did not occur between stage 2 and combined stages 3 and 4.

Since the depth of sleep typically produces measurable changes in autonomic bodily functions, we anticipated that autonomic activity could be used to describe depth of sleep. Of the organs under the control of the autonomic nervous system, the heart has one of the richest supplies of both adrenergic and cholinergic nerve endings. The heart also produces an electrical signal that is easy to measure (the electrocardiogram). Most of the information supplied to the heart by the autonomic nervous system is reflected in the instantaneous R to R interval (the beat-by-beat heart rate). Only a small portion of the autonomic information supplied to the heart is reflected in the electrocardiographic wave shape. Thus the beat-by-beat heart rate should be a useful measure in determining sleep stages.

In a study by Brooks ( 6 ) six individuals (three husbands and wives) were observed for fifty nights of sleep, Brooks found a 10% average increase in heart rate when the depth of sleep lessened by one stage from stages 4 to 3 ot 3 to 2. He also found a 13.7% average increase in heart rate with two stage lessening of sleep (4 to 2 and 3 to 1). A 21.5% average increase in heart rate occurred when sleep level lightened by three stages (from 4 to 1 or from stage 3 to wakefulness). Brooks concluded that sleep depth was probably reflected more in changes in cardiac cycle length (i.e. instantaneous R to R interval or beat-by-beat heart rate) than in the average heart rate values he used.

## Review of Our Past Work

In our previous work, analysis of variance was applied to several types of measurements which were obtained from instantaneous heart rate data. Significant measures of sample mean value, sample standard deviation and histogram were found for both the instantaneous heart rate and the beat-by-beat interval. Table 1 graphically illustrates the level of significance of each of these measures for each subject for the interval measures. The level of significance is measured with the F test and indicates the level of rejection of the hypotheses of equal mean value for each stage of sleep for the measure under consideration.

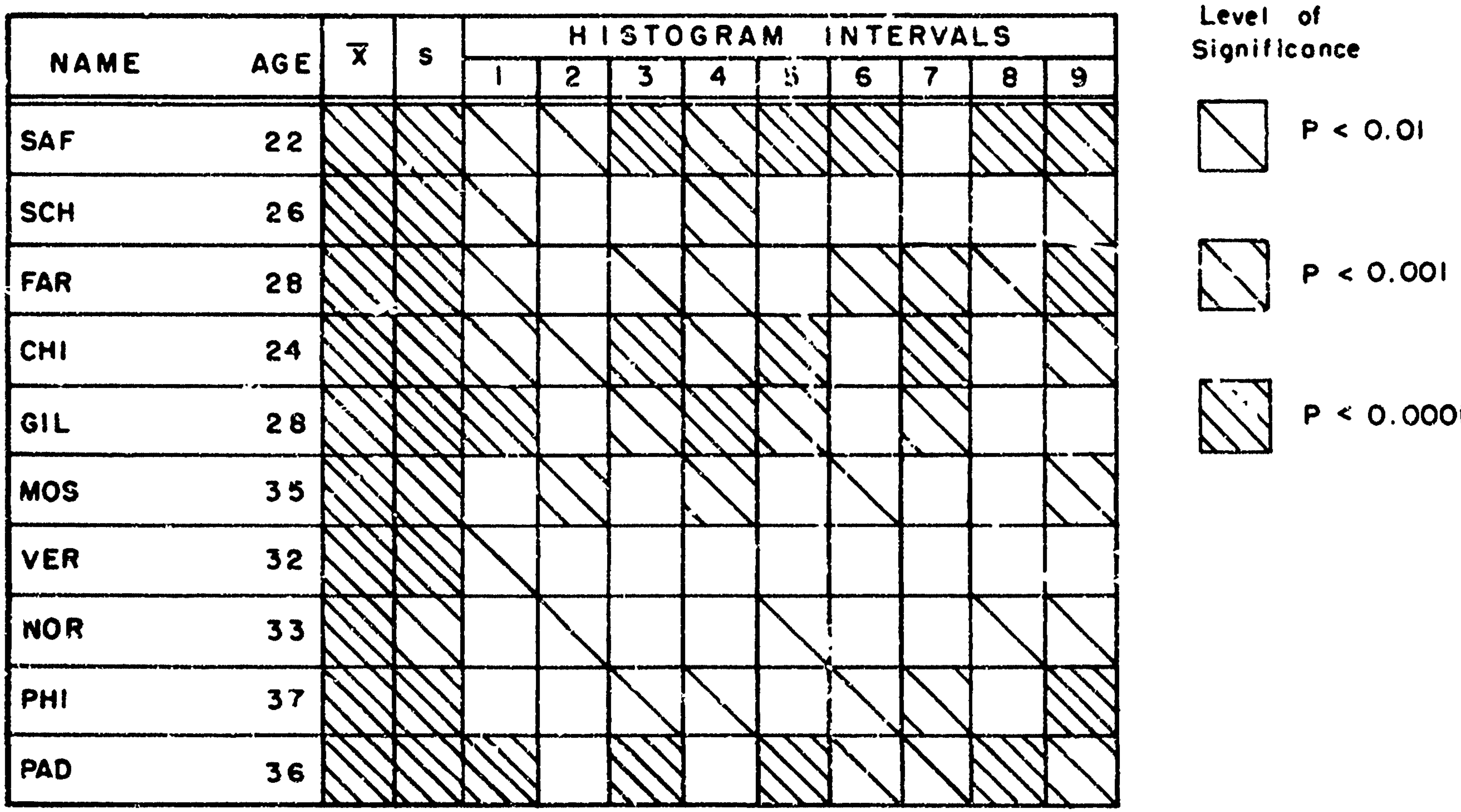
Fourier analysis of the instantaneous heart rate data produced a large number of variables which were significant at the .001 level. All measures were made on ensembles of 128 heart beat intervals (the number of seconds between each heart beat). A detailed presentation of the results of this work is available (21). At the time we wrote this earlier report, it was recognized that the analysis of variance test did not define the level of separation a variable might accomplish in multiple class data. However, variables that are significant at the .001 level can frequently provide a reasonable starting point in the search for reliable measures to be used for classification.

## Rationale of Present Approach

The data of all test subjects presented to The University of Texas was classified into one of six states (or levels) of consciousness. These were

TABLE 1

SIGNIFICANT BEAT-TO-BEAT INTERVAL DESCRIPTORS



awake, stage one through four sleep, and stage REM. While these categories or stages of consciousness have been considered by many to be standard, they are based primarily on EEG criteria. Physiologically these stages probably have little direct meaning.

Because of the difficulty of computer classifying six categories, we attempted to simplify the problem. The deep sleep stages 3 and 4 were combined into a single stage and stage 1 and REM were grouped together. Therefore the number of consciousness levels was reduced from 6 to 4.

Stage 1 and REM were combined because we felt there was little meaningful difference between the two sleep stages. Sleep stage REM is usually scored whenever Stage 1 EEG is found after the 1st sleep cycle. In other words, the only time Stage 1 was scored was during the first sleep cycle; after that all Stage 1 EEG was typically scored as REM. Agnew (2) concurred in this opinion.

Similarly, we were unable to find any reason why Stages 3 and 4 should not be combined. Typically in the evaluation of sleep effectiveness, stages 3 and 4 are combined (2).

Once the dimensionality of our problem was reduced to four, we sought to develop a "cost" function to evaluate classification procedure. This was necessary since it is typically difficult for humans to compare error functions of four variables. The weightings suggested by Agnew (2) stress the importance of differentiating between awake and sleep.

A much smaller cost is assigned for making "one" stage errors during sleep. Table 2 represents the empirically-derived cost matrix.

TABLE 2

The cost of classifying stage - - -

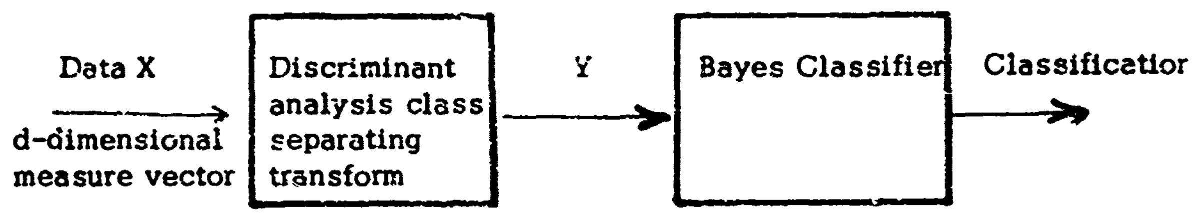
| ...as stage | | Awake | 1, REM | 2 | 3,4 |
|---|---|---|---|---|---|
| | Awake | 0 | 1 | 1 | 1 |
| | 1,REM | 1 | 0 | 1/4 | 1/2 |
| | 2 | 1 | 1/4 | 0 | 1/4 |
| | 3,4 | 1 | 1/2 | 1/4 | 0 |

A cost per classification is obtained when this cost matrix is multiplied by a sleep result matrix (or an element-by-element basis - not as matrix multiplication is typically performed) and the elements of this product matrix are summed then divided by the total number of points considered in the classification matrix.

Note that the greatest cost of errors occurs when an awake epoch is classified as sleep, or when any sleep epoch is classified as awake. No cost is accrued for a correct answer and weightings of one-half are charged for missing the sleep stage by more than one level. This single number cost value permits effective comparison of the various variables used in the classification procedure.

## DISCRIMINANT ANALYSIS

The optimum procedure for classifying a sample of data with d-measures is achieved using the Bayes discriminant function which requires knowledge (or estimates) of the *a priori* probability of occurrence of each category and the d-dimensional joint density function of the measures for each category. A non-optimal, but computationally feasible, approach to classification is Multiple Discriminant Analysis. This analysis includes a linear transformation to reduce the dimensionality of the problem and a Bayes classifier as illustrated in Figure 1




CLASSIFICATION PROCEDURE

Figure 1

Often information from an experiment can be divided into a sequence of consecutive epochs. For each epoch a set of descriptors or measures is computed that may contain sufficient information for the classification of the epoch.

It is convenient to picture the set of data for each epoch as a point in an d-dimensional space where one point describes each epoch. Further assume the classification associated with each of these epochs is known; where

$i$ is equal to 1, 2, 3,...R and R is the total number of classes.

Another important representation of the d descriptors is that of a d-dimensional vector X. Each epoch is represented by a different vector. Thus we have the picture of points in a d-dimensional space and their corresponding vector representation X.

The linear transformation of Figure 1 maximizes the distance between centroids of each category in the Y space while holding the overall variance constant. The transformation reduces the dimensionality to the minimum number required to compartmentize the space for the categories under consideration. That is, the dimensionality of Y is the minimum of either (a) dimensions of X, or (b) the number of categories minus one.

The discriminant analysis procedure assumes the joint density function of Y for each category is normally distributed. Thus, conditional probability of group occurrence may be computed according to:

$$P(i/Y) = \frac{\pi_i f_i}{\sum_{j=1}^{R} \pi_j f_j} \qquad i = 1, 2, \ldots R$$

where

$\pi_j$ = is the a priori probability of occurrence of class j

$f_j$ is the joint density function for class j evaluated at Y.

$$F_j = \frac{1}{(2\pi)^{d/2} |\Sigma_j|^{\frac{1}{2}}} e^{-\frac{1}{2}(Y-\mu_j)' \Sigma_j^{-1} (Y-\mu_j)}$$

$\mu_j$ is the centroid of class j in the Y space

$\Sigma_j$ is the covariance matrix of Y for class j

Application to Heart Rate Data

The heart rate data was divided into a sequence of 128 beat-to-beat epochs. For each epoch heart rate descriptors were computed that were anticipated to contain sufficient information for the determination of sleep stage during that epoch.

# PROCEDURE

## Data Acquisition

Electroencephalographic and electrocardiographic data for this study were collected on FM magnetic tape at The University of Florida Sleep Laboratory by W. H. Agnew, Jr. under Air Force Contract No. F41609-68-C-003. Ten subjects at The University of Florida were selected on the basis of good physical and mental health as determined by medical examination and the Minnesota Multiphasic Personality Inventory. Each subject spent at least three consecutive nights in the laboratory. The first night was used to condition the subjects to the laboratory in order to avoid first night effects ( 1 ). Eight hours of sleep EEG, EOG, and electrocardiograms were recorded on FM magnetic tape for two nights. The method of data collection and sleep scoring of the EEG was reported in detail by Lessard, Ford and Hughes of the USAF School of Aerospace Medicine (14). Copies of the FM magnetic tapes containing the sleep data were supplied to The University of Texas Bio-Medical Engineering Laboratory by the U. S. Air Force School of Aerospace Medicine.

## Data Reduction

Double differentiation of the filtered electrocardiogram and threshold logic units were used to detect the peak of each R-wave. The time in milliseconds between heart beats was written on digital magnetic tape. The data were grouped into records containing 128 consecutive beat-to-beat intervals.

A number of descriptors were then computed for each record. Eleven of the descriptors were the mean value $\bar{X}$, the sample variance $S^2$, and the nine intervals histogram of standardized instantaneous beat-to-beat intervals for each record. Standardized beat-to-beat intervals were calculated using equation (1).

$$Z_i = \frac{X_i - \bar{X}}{S} \tag{1}$$

The R-R interval for each beat was $X_i$ and the mean value and standard deviation of the 128 R-R intervals was $\bar{X}$ and S, respectively. The standardized instantaneous value for the ith heart beat is $Z_i$. The distribution of Z scores calculated in this way always has a mean value of zero and standard deviation of one. Other measurements for the 128 R-R intervals included eleven instantaneous heart rate measures analagous to the interval measures and 64 Fourier Transform measures.

Histogram intervals of one-half standard deviation were selected as illustrated in Table 3. The number of standardized scores falling in each of the nine intervals were used as the value of the histogram descriptors. Each group of 128 heart intervals had a different mean value, $\bar{X}$ and sample standard deviation, S. The outer intervals of the histogram extended from $-\infty$ to $-$ 1-3/4 and $+$ 1-3/4 to $+\infty$.

Analysis of variance (7,25) was used to determine measure significance for each subject. This procedure tested the null hypothesis for each measure. In other words, what is the probability that a measure's mean value is the

TABLE 3

HEART INTERVAL MEASURES

| Measures | Description of the Measure |
|---|---|
| 1 | Sample Mean Value |
| 2 | Sample Standard Deviation |
| | Histogram Measures (1/2 $\sigma$ Intervals) |
| 3 | $x < -1\,3/4\,\sigma$ |
| 4 | $-1\,3/4\,\sigma \leq x < -1\,1/4\,\sigma$ |
| 5 | $-1\,1/4\,\sigma \leq x < -3/4\,\sigma$ |
| 6 | $-3/4\,\sigma \leq x < -1/4\,\sigma$ |
| 7 | $-1/4\,\sigma \leq x < -1/4\,\sigma$ |
| 8 | $1/4\,\sigma \leq x < 3/4\,\sigma$ |
| 9 | $3/4\,\sigma \leq x < 1\,1/4\,\sigma$ |
| 10 | $1\,1/4\,\sigma \leq x < 1\,3/4\,\sigma$ |
| 11 | $1\,3/4\,\sigma \leq x$ |

same for awake and the five stages of sleep. Our program used an F-test to compute the probability that the mean values of the descriptors for each stage of sleep were equal. The hypothesis was rejected if $P < .01$. The most significant measures were selected to be used to train the discriminant function. Detailed results of the analysis of variance procedure are in our previous report (21).

Discriminant Analysis

After descriptors for each epoch of 128 heart beats had been computed, computer classification was performed in the following manner.

(1) Approximately 50% of the first night's epochs were selected as training data. The exact number of epochs corresponded to the larger of either 25% of that class's total number of epochs for an individual's two nights of sleep or ten epochs. Ten epochs represented an arbitrary minimum number of samples for estimation of the d-dimensional centroid and covariance matrix for each class. If the first night of sleep did not contain the required number of epochs then epochs were selected from the second night of sleep. Each epoch contained a complete set of heart rate, interval and Fourier transform descriptors.

(2) Selected subsets of descriptors from the training set were entered into the discriminant analysis program which evaluated the class separating linear transformation $X \rightarrow Y$ and computed mean values and covariance matrices used in the Baysean conditional probability estimates.

(3) The accuracy of selected subset of descriptors was pretested by computing the conditional probability P(i/X) for each epoch in the training set and comparing the results to the EEG scored sleep stages.

(4) The two nights of recorded sleep then were classified by transforming the selected subset of descriptors one epoch at a time and computing the conditional probability of classification for each epoch. The computer classifications were recorded on magnetic tape for display and accuracy computations.

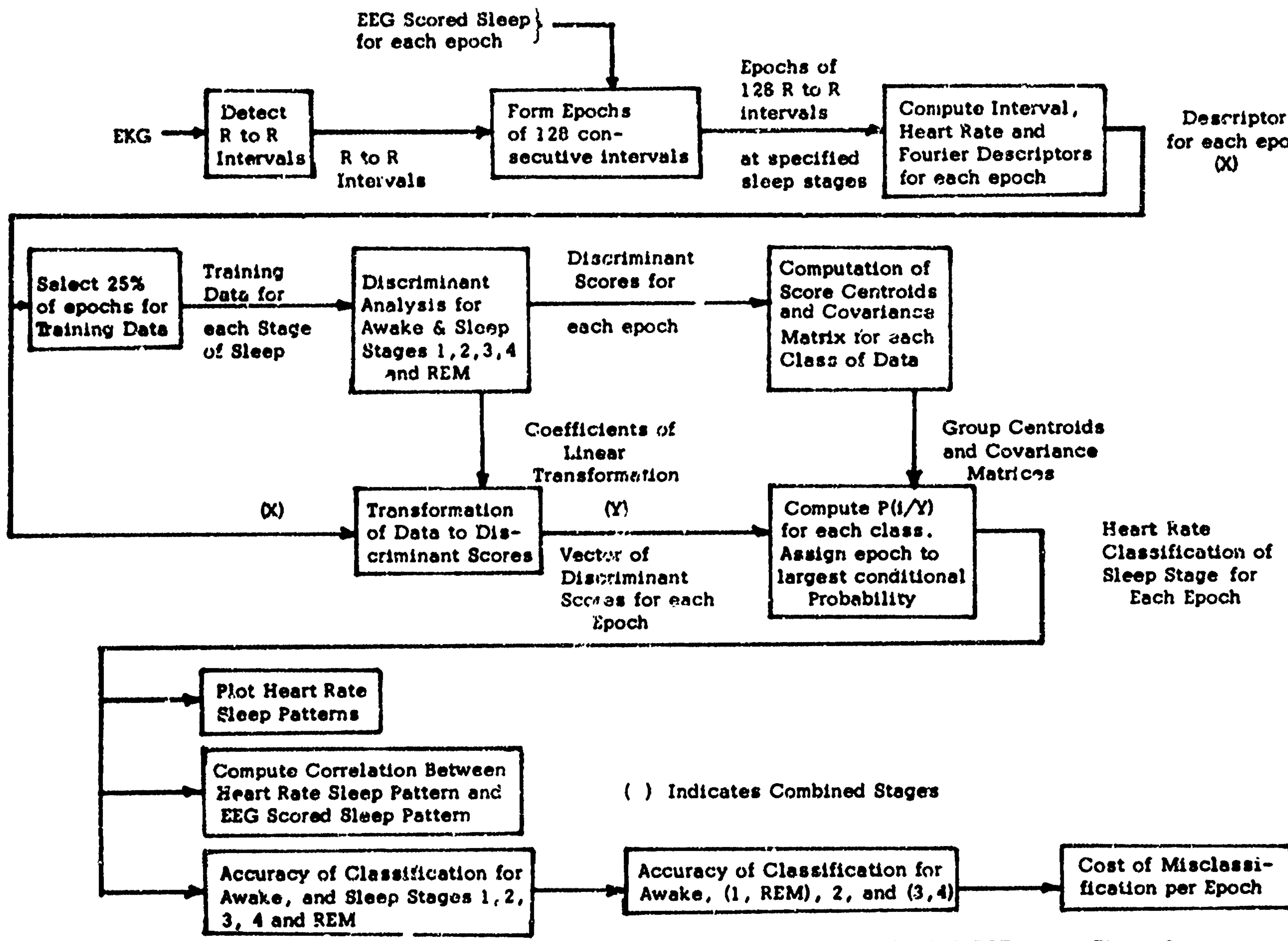
(5) The computer classified sleep patterns were plotted on a CalComp plotter and the accuracy was determined by

(a) presentation of an error table and computation of the percent correctness

(b) computation of the correlation coefficient between the computer scored and EEG scored sleep records

(c) pooling the data into four classes (i) awake, (ii) 1, 1-REM, (iii) 2, and (iv) 3, 4 and evaluating the average cost per epoch based upon cost function presented in Table 2, page 7.

A flow graph of the analysis procedure is shown in Figure 2.




ANALYSIS PROCEDURE FOR EACH SUBJECT - Figure 2

# RESULTS

The results of this study are presented in both tabular and graphic form. Tables 4 through 22 present the tabular data. The even-numbered tables (4, 6, 8, 10, etc.) contain the results of classifying the first night of recorded sleep and the odd-numbered tables contain the results for the second night of recorded sleep. In most cases, the discriminant algorithm was trained on data obtained from the first night of recorded sleep. Results are presented for using both balanced (all classes equally likely) and unbalanced (actual frequency of occurrence of each category) *a priori* probabilities. For the specified sets of descriptors the tables contain

(1) Accuracy for both six and four category classifications

(2) Average cost per epoch for the night

(3) Correlation coefficient between EEG hand-scored sleep heart rate computer-scored sleep.

In these tables, Variable 1 is the mean beat-to-beat interval of the epoch, Variable 2 is the sample variance of each epoch, and Variable 3 through 11 are the interval histogram values. Table 3, summarizes the relation between variable number and physical beat-to-beat interval measures. (For any epoch the sum of the histogram measures is 128).

The variables noted in the tables as "11 Fourier" represent 11 of the best Fourier variables selected by analysis of variance (21). The measure set "Histogram and Four" combines the 8 Interval Histogram measures with the

11 Fourier measures.

The conditional probabilities of each sleep stage were calculated assuming each sleep stage to be equally likely (balanced) and using a priori probabilities actually based on the individual subject. The a priori probabilities of each of the sleep levels for each subject is presented in Table 24.

Five measures of merit of the classification procedure are presented for each combination of measures used and for each set of a priori probabilities. The percent classification of the first half and the second half of each night using six sleep categories are presented. In addition, the percent correct classification using four categories is presented. The weighted cost per epic is listed as a measure of machine scoring effectiveness. The correlation coefficient given is a rough indication of the sameness in shape between the hand-scored and machine-scored data.

Tables 5 through 23 and all odd-numbered tables in between represent the results of discriminant classification using a discriminant function trained on the first night of sleep and used to classify the second night of sleep. The same variables and a priori probabilities presented in the preceding even-numbered table are used in the odd-numbered table. The measures of merit are the same as those used for the preceding even-numbered tables. Occasionally when insufficient samples were available in the first night of data, a few samples had to be obtained from the second night of sleep to train the discriminant function.

TABLE 4

CLASSIFICATION OF THE FIRST NIGHT OF RECORDED SLEEP BY A DISCRIMINANT FUNCTION

| Subject | Measures used | A Priori Probability | % Correct File 1* | % Correct File 2** | % Correct Both Files | Cost Per Epoch | Correlation† |
|---|---|---|---|---|---|---|---|
| Safer | 2,4-11 | Not Balanced | 64.44 | 36.07 | 56.07 | 0.193 | 0.676 |
| Safer | 2,4-11 | Balanced | 61.46 | 31.97 | 53.33 | 0.192 | 0.667 |
| Safer | 2,4-10 | Not Balanced | 61.48 | 39.34 | 55.68 | 0.165 | 0.669 |
| Safer | 2,4-10 | Balanced | 60.0 | 31.15 | 54.11 | 0.180 | 0.703 |
| Safer | 1,4-10 | Not Balanced | 68.15 | 41.8 | 60.0 | 0.132 | 0.684 |
| Safer | 1,4-10 | Balanced | 63.7 | 35.25 | 58.03 | 0.131 | 0.693 |

*File 1 was the first half of the first night of data on each subject. Most of the training set came from file 1.

**File 2 was the second half of the first night of data on each subject. Some of the training set came from this data file.

†Correlation between entire night of sleep as machine scored against the hand scored night of sleep.

TABLE 5

CLASSIFICATION OF THE SECOND NIGHT OF RECORDED

SLEEP WITH DISCRIMINANT ANALYSIS USING TRAINING DATA FROM THE FIRST NIGHT OF SLEEP

| Subject | Variables | A Priori Probability | 6 Categories % Correct File 3* | 6 Categories % Correct File 4** | 4 Categories % Correct Both Files | Cost Per Epoch | Correlation† |
|---|---|---|---|---|---|---|---|
| Safer | 2.4-11 | Not Balanced | 26.62 | 37.90 | 32.95 | 0.445 | 0.131 |
| Safer | 2,4-11 | Balanced | 26.62 | 29.84 | 29.8 | 0.462 | 0.122 |
| Safer | 2,4-10 | Not Balanced | 23.74 | 44.35 | 34.48 | 0.449 | 0.121 |
| Safer | 2,4-10 | Balanced | 28.78 | 29.03 | 30.65 | 0.444 | 0.161 |
| Safer | 1,4-10 | Not Balanced | 33.09 | 34.68 | 36.39 | 0.412 | 0.357 |
| Safer | 1,4-10 | Balanced | 35.25 | 29.03 | 36.01 | 0.399 | 0.375 |

*roughly the first half of the second night of data was on file 3 for each subject

**the second half of the second night's sleep was on file 4.

†correlation between the entire night of hand scored sleep against the entire night of machine scored sleep.

TABLE 6

CLASSIFICATION OF THE FIRST NIGHT OF RECORDED SLEEP BY A DISCRIMINANT FUNCTION

| Subject | Measures Used | A Priori Probability | % Correct File 1* | % Correct File 2** | % Correct Both Files | Cost Per Epoch | Correlation† |
|---|---|---|---|---|---|---|---|
| Farrington | 11 Fourier | Not Balanced | 59.57 | 50.0 | 58.06 | 0.149 | 0.408 |
| Farrington | 11 Fourier | Balanced | 39.36 | 32.61 | 44.62 | 0.239 | 0.393 |
| Chinoy | Hist & Four | Not Balanced | 66.67 | 67.39 | 71.87 | 0.105 | [illegible] |
| Chinoy | Hist & Four | Balanced | 66.67 | 63.04 | [illegible] | 0.104 | 0.675 |
| Chinoy | 11 Fourier | Not Balanced | 36.27 | 33.69 | 42.19 | 0.238 | 0.228 |
| Chinoy | 11 Fourier | Balanced | 35.29 | 21.74 | 39.06 | 0.266 | 0.237 |

*File 1 was the first half of the first night of data on each subject. Most of the training set came from file 1.

**File 2 was the second half of the first night of data on each subject. Some of the training set came from this data file.

†Correlation between entire night of sleep as machine scored against the hand scored night of sleep.

TABLE 7

CLASSIFICATION OF THE SECOND NIGHT OF RECORDED

SLEEP WITH DISCRIMINANT ANALYSIS USING TRAINING DATA FROM THE FIRST NIGHT OF SLEEP

| | | | 6 Categories | | 4 Categories | | |
|---|---|---|---|---|---|---|---|
| Subject | Variables | A Priori Probabilities | % Correct File 3* | % Correct File 4** | % Correct Both Files | Cost Per Epoch | Correlation† |
| Farrington | 11 Fourier | Not Balanced | 47.91 | 49.46 | 55.03 | 0.169 | 0.488 |
| Farrington | 11 Fourier | Balanced | 41.67 | 26.88 | 44.97 | 0.242 | 0.531 |
| Chinoy | Hist & Four | Not Balanced | 52.68 | 58.16 | 65.75 | 0.116 | 0.473 |
| Chinoy | Hist & Four | Balanced | 44.01 | 48.98 | 59.11 | 0.137 | 0.477 |
| Chinoy | 11 Fourier | Not Balanced | 17.20 | 41.84 | 36.46 | 0.251 | -0.044 |
| Chinoy | 11 Fourier | Balanced | 9.67 | 28.57 | 30.38 | 0.291 | 0.047 |

*roughly the first half of the second night of data was on file 3 for each subject.

**the second half of the second night's sleep was on file 4.

†correlation between the entire night of hand scored sleep against the entire night of machine scored sleep.

TABLE 8

CLASSIFICATION OF THE FIRST NIGHT OF RECORDED SLEEP BY A DISCRIMINANT FUNCTION

| Subject | Measures Used | A Priori Probabilities | % Correct File 1* | % Correct File 2** | % Correct Both Files | Cost Per Epoch | Correlation† |
|---|---|---|---|---|---|---|---|
| Chinoy | 1,4-11 | Not Balanced | 67.96 | 71.72 | 73.5 | 0.108 | 0.640 |
| Chinoy | 1,4-11 | Balanced | 61.17 | 65.66 | 68.0 | 0.128 | 0.610 |
| Chinoy | 1,2,3,5, 7,9,11 | Not Balanced | 71.84 | 73.74 | 77.5 | 0.084 | 0.725 |
| Chinoy | 1,2,3,5, 7,9,11 | Balanced | 70.87 | 67.68 | 75.0 | 0.091 | 0.714 |
| Gildersleeve | 1,4-11 | Not Balanced | 58.33 | 60.75 | 63.11 | 0.149 | 0.436 |
| Gildersleeve | 1.4-11 | Balanced | 52.50 | 46.73 | 56.44 | 0.170 | 0.394 |

*File 1 was the first half of the first night of data on each subject. Most of the training set came from file 1.

**File 2 was the second half of the first night of data on each subject. Some of the training set came from this data file.

†Correlation between entire night of sleep as machine scored against the hand scored night of sleep.

TABLE 9

CLASSIFICATION OF THE SECOND NIGHT OF RECORDED

SLEEP WITH DISCRIMINANT ANALYSIS USING TRAINING DATA FROM THE FIRST NIGHT OF SLEEP

| | | | 6 Categories | | 4 Categories | | |
|---|---|---|---|---|---|---|---|
| Subject | Variables | A Priori Probabilities | % Correct File 3* | % Correct File 4** | % Correct Both Files | Cost Per Epoch | Correlation† |
| Chinoy | 1,4-11 | Not Balanced | 64.52 | 52.88 | 65.77 | 0.120 | 0.514 |
| Chinoy | 1,4-11 | Balanced | 56 9 | 47.1 | 62.03 | 0.14 | 0.464 |
| Chinoy | 1,2,3,5, 7,9,11 | Not Balanced | 52.69 | 60.58 | 66.84 | 0.120 | 0.557 |
| Chinoy | 1,2,3,5, 7,9,11 | Balanced | 48.39 | 50.96 | 63.10 | 0.132 | 0.556 |
| Gildersleeve | 1,4-11 | Not Balanced | 39.52 | 60.91 | 56.09 | 0.156 | 0.316 |
| Gildersleeve | 1,4-11 | Balanced | 36.29 | 59.09 | 54.35 | 0.175 | 0.308 |

*roughly the first half of the second night of data was on file 3 for each subject.

**the second half of the second night's sleep was on file 4.

†correlation between the entire night of hand scored sleep against the entire night of machine scored sleep.

TABLE 10

CLASSIFICATION OF THE FIRST NIGHT OF RECORDED SLEEP BY A DISCRIMINANT FUNCTION

| Subject | Measures Used | A Priori Probabilities | % Correct File 1* | % Correct File 2** | % Correct Both Files | Cost Per Epoch | Correlation† |
|---|---|---|---|---|---|---|---|
| Farrington | 1,4-10 | Not Balanced | 62.11 | 46.24 | 59.04 | 0.137 | 0.468 |
| Farrington | 1,4-10 | Balanced | 47.37 | 41.94 | 49.43 | 0.171 | 0.420 |
| Farrington | 2,4-10 | Not Balanced | 56.84 | 54.84 | 57.97 | 0.164 | 0.345 |
| Farrington | 2,4-10 | Balanced | 41.05 | 40.86 | 46.8 | 0.234 | 0.403 |
| Safer | 4-11 | Not Balanced | 55.56 | 38.52 | 51.7 | 0.207 | 0.599 |
| Safer | 4-11 | Balanced | 51.11 | 34.34 | 50.58 | 0.207 | 0.635 |
| Safer | 1,4-11 | Not Balanced | 65.93 | 46.72 | 61.18 | 0.132 | 0.696 |
| Safer | 1,4-11 | Balanced | 64.44 | 36.89 | 57.64 | 0.137 | 0.691 |

*File 1 was the first half of the first night of data on each subject. Most of the training set came from file 1.

**File 2 was the second half of the first night of data on each subject. Some of the training set came from this data file.

†Correlation between entire night of sleep as machine scored against the hand scored night of sleep.

TABLE 11

CLASSIFICATION OF THE SECOND NIGHT OF RECORDED

SLEEP WITH DISCRIMINANT ANALYSIS USING TRAINING DATA FROM THE FIRST NIGHT OF SLEEP

| | | | 6 Categories | | 4 Categories | | |
|---|---|---|---|---|---|---|---|
| Subject | Variables | A Priori Probabilities | % Correct File 3* | % Correct File 4** | % Correct Both Files | Cost Per Epoch | Correlation† |
| Farrington | 1,4-10 | Not Balanced | 58.7 | 56.3 | 64.39 | 0.130 | 0.639 |
| Farrington | 1,4-10 | Balanced | 53.61 | 44.68 | 59.16 | 0.145 | 0.652 |
| Farrington | 2,4-10 | Not Balanced | 44.33 | 61.70 | 57.07 | 0.212 | 0.42 |
| Farrington | 2,4-10 | Balanced | 44.33 | 48.93 | 52.88 | 0.240 | 0.498 |
| Safer | 4-11 | Not Balanced | 26.62 | 25.81 | 27.2 | 0.490 | 0.121 |
| Safer | 4-11 | Balanced | 26.62 | 20.16 | 26.05 | 0.489 | 0.130 |
| Safer | 1,4-11 | Not Balanced | 35.25 | 34.68 | 37.16 | 0.397 | 0.343 |
| Safer | 1,4-11 | Balanced | 33.09 | 28.23 | 34.10 | 0.410 | 0.359 |

*roughly the first half of the second night of data was on file 3 for each subject

**the second half of the second night's sleep was on file 4.

†correlation between the entire night of hand scored sleep against the entire night of machine scored sleep.

TABLE 12

CLASSIFICATION OF THE FIRST NIGHT OF RECORDED SLEEP BY A DISCRIMINANT FUNCTION

| Subject | Measures Used | A Priori Probabilities | % Correct File 1* | % Correct File 2** | % Correct Both Files | Cost Per Epoch | Correlation† |
|---|---|---|---|---|---|---|---|
| Schmidt | 1,4-11 | Not Balanced | 67.29 | 32.32 | 55.3 | 0.147 | 0.472 |
| Schmidt | 1,4-11 | Balanced | 57.01 | 23.23 | 46.6 | 0.192 | 0.436 |
| Schmidt | 2,4-11 | Not Balanced | 52.34 | 21.05 | 59.7 | 0.160 | 0.346 |
| Schmidt | 2,4-11 | Balanced | 42.99 | 40.40 | 49.0 | 0.279 | 0.273 |
| Schmidt | 2,4-10 | Not Balanced | 57.94 | 54.55 | 61.83 | 0.166 | 0.35 |
| Schmidt | 2,4-10 | Balanced | 50.47 | 42.42 | 52.9 | 0.243 | 0.232 |
| Schmidt | 1,4-10 | Not Balanced | 65.4 | 36.4 | 55.8 | 0.148 | 0.476 |
| Schmidt | 1,4-10 | Balanced | 58.88 | 25.25 | 50.0 | 0.172 | 0.474 |

*File 1 was the first half of the first night of data on each subject. Most of the training set came from file 1.

**File 2 was the second half of the first night of data on each subject. Some of the training set came from this data file.

†Correlation between entire night of sleep as machine scored against the hand scored night of sleep.

TABLE 13

CLASSIFICATION OF THE SECOND NIGHT OF RECORDED

SLEEP WITH DISCRIMINANT ANALYSIS USING TRAINING DATA FROM THE FIRST NIGHT OF SLEEP

| Subject | Variables | A Priori Probabilities | 6 Categories % Correct File 3* | 6 Categories % Correct File 4** | 4 Categories % Correct Both Files | Cost Per Epoch | Correlation † |
|---|---|---|---|---|---|---|---|
| Schmidt | 1,4-11 | Not Balanced | 17.53 | 18.48 | 34.6 | 0.282 | 0.27 |
| Schmidt | 1,4-11 | Balanced | 16.49 | 17.39 | 30.3 | 0.301 | 0.254 |
| Schmidt | 2,4-11 | Not Balanced | 21.65 | 33.70 | 45.4 | 0.25 | 0.26 |
| Schmidt | 2,4-11 | Balanced | 24.74 | 23.91 | 43.24 | 0.266 | 0.314 |
| Schmidt | 2,4-10 | Not Balanced | 24.74 | 34.78 | 44.38 | 0.255 | 0.262 |
| Schmidt | 2,4-10 | Balanced | 29.9 | 25.0 | 43.7 | 0.271 | 0.287 |
| Schmidt | 1,4-10 | Not Balanced | 16.49 | 20.65 | 34.05 | 0.287 | 0.208 |
| Schmidt | 1,4-10 | Balanced | 16.49 | 18.48 | 33.5 | 0.289 | 0.240 |

*roughly the first half of the second night of data was on file 3 for each subject

**the second half of the second night's sleep was on file 4.

†correlation between the entire night of hand scored sleep against the entire night of machine scored sleep.

TABLE 14

CLASSIFICATION OF THE FIRST NIGHT OF RECORDED SLEEP BY A DISCRIMINANT FUNCTION

| Subject | Measures Used | A Priori Probabilities | % Correct File 1* | % Correct File 2** | % Correct Both Files | Cost Per Epoch | Correlation † |
|---|---|---|---|---|---|---|---|
| Nordyke | 1,4-11 | Not Balanced | 50.48 | 48.42 | 55.10 | 0.1696 | 0.425 |
| Nordyke | 1,4-11 | Balanced | 42.86 | 40.00 | 47.95 | 0.213 | 0.3898 |
| Phillips | 1,4-11 | Not Balanced | 60.19 | 35.92 | 50.00 | 0.262 | 0.419 |
| Phillips | 1,4-11 | Balanced | 48.15 | 30.10 | 42.31 | 0.284 | 0.493 |
| Padula | 1,4-11 | Not Balanced | 57.55 | 36.29 | 52.895 | 0.249 | 0.233 |
| Padula | 1,4-11 | Balanced | 43.17 | 39.52 | 46.72 | 0.295 | 0.264 |
| Moss | 1,4-11 | Not Balanced | 60.32 | 44.9 | 58.93 | 0.2098 | 0.362 |
| Moss | 1,4-11 | Balanced | 52.38 | 40.68 | 53.57 | 0.272 | 0.390 |

*File 1 was the first half of the first night of data on each subject. Most of the training set came from file 1.

**File 2 was the second half of the first night of data on each subject. Some of the training set came from this data file.

†Correlation between entire night of sleep as machine scored against the hand scored night of sleep.

TABLE 15

CLASSIFICATION OF THE SECOND NIGHT OF RECORDED

SLEEP WITH DISCRIMINANT ANALYSIS USING TRAINING DATA FROM THE FIRST NIGHT OF SLEEP

| Subject | Variables | A Priori Probabilities | 6 Categories % Correct File 3* | 6 Categories % Correct File 4** | 4 Categories % Correct Both Files | Cost Per Epoch | Correlation† |
|---|---|---|---|---|---|---|---|
| Nordyke | 1,4-11 | Not Balanced | 37.27 | 31.07 | 35.89 | 0.300 | 0.182 |
| Nordyke | 1,4-11 | Balanced | 30.91 | 27.18 | 33.49 | 0.331 | 0.140 |
| Phillips | 1,4-11 | Not Balanced | 42.47 | 47.89 | 49.31 | 0.213 | 0.176 |
| Phillips | 1,4-11 | Balanced | 32.88 | 40.85 | 43.06 | 0.257 | 0.251 |
| Padula | 1,4-11 | Not Balanced | 23.58 | 14.41 | 23.95 | 0.503 | 0.086 |
| Padula | 1,4-11 | Balanced | 21.14 | 11.86 | 22.69 | 0.526 | 0.110 |
| Moss | 1,4-11 | Not Balanced | 21.48 | 29.84 | 34.94 | 0.353 | 0.0357 |
| Moss | 1,4-11 | Balanced | 11.11 | 25.0 | 28.51 | 0.416 | 0.027 |

*roughly the first half of the second night of data was on file 3 for each subject.

**the second half of the second night's sleep was on file 4.

†correlation between the entire night of hand scored sleep against the entire night of machine scored sleep.

TABLE 16

CLASSIFICATION OF THE FIRST NIGHT OF RECORDED SLEEP BY A DISCRIMINANT FUNCTION

| Subject | Measures Used | A Priori Probabilities | % Correct File 1* | % Correct File 2** | % Correct Both Files | Cost Per Epoch | Correlation† |
|---|---|---|---|---|---|---|---|
| Schmidt | Hist & Four | Not Balanced | 47.25 | 34.38 | 44.39 | 0.185 | 0.323 |
| Schmidt | Hist & Four | Balanced | 45.06 | 28.13 | 40.1 | 0.201 | 0.304 |
| Schmidt | $^{11}$Four | Not Balanced | 29.67 | 42.71 | 36.89 | 0.290 | -0.236 |
| Schmidt | $^{11}$Four | Balanced | 18.68 | 27.08 | 24.59 | 0.348 | -0.233 |

*File 1 was the first half of the first night of data on each subject. Most of the training set came from file 1.

**File 2 was the second half of the first night of data on each subject. Some of the training set came from this data file.

†Correlation between entire night of sleep as machine scored against the hand scored night of sleep.

TABLE 17

CLASSIFICATION OF SECOND NIGHT OF RECORDED

SLEEP WITH DISCRIMINANT ANALYSIS USING TRAINING DATA FROM THE FIRST NIGHT OF SLEEP

| | | | 6 Categories | | 4 Categories | | |
|---|---|---|---|---|---|---|---|
| Subject | Variables | A Priori Probabilities | % Correct File 3* | % Correct File 4** | % Correct Both Files | Cost Per Epoch | Correlation† |
| Schmidt | Hist & Four | Not Balanced | 12.37 | 19.56 | 34.05 | 0.279 | 0.443 |
| Schmidt | Hist & Four | Balanced | 15.46 | 18.48 | 34.05 | 0.279 | 0.334 |
| Schmidt | $^{11}$Four | Not Balanced | 21.35 | 29.34 | 29.72 | 0.319 | -0.084 |
| Schmidt | $^{11}$Four | Balanced | 15.46 | 19.57 | 25.95 | 0.362 | -0.091 |

*roughly the first half of the second night of data was on file 3 for each subject.

**the second half of the second night's sleep was on file 4.

†correlation between the entire night of hand scored sleep against the entire night of machine scored sleep.

TABLE 18

CLASSIFICATION OF THE FIRST NIGHT OF RECORDED

SLEEP BY A DISCRIMINANT FUNCTION

| Subject | Measures Used | A Priori Probabilities | % Correct File 1* | % Correct File 2** | % Correct Both Files | Cost Per Epoch | Correlation† |
|---|---|---|---|---|---|---|---|
| Phillips | Hist 1,4-11 & $^{11}$Fourier | Not Balanced | 64.49 | 36.27 | 55.56 | 0.245 | 0.484 |
| Phillips | Hist 1,4-11 & $^{11}$Fourier | Balanced | 51.40 | 39.39 | 46.86 | 0.268 | 0.481 |
| Phillips | $^{11}$Fourier | Not Balanced | 58.88 | 41.18 | 53.14 | 0.307 | 0.384 |
| Phillips | $^{11}$Fourier | Balanced | 40.19 | 25.49 | 38.64 | 0.388 | 0.332 |
| Moss | Hist 1,4-11 & $^{11}$Four | Not Balanced | 65.60 | 49.57 | 66.67 | 0.189 | 0.349 |
| Moss | Hist 1,4-11 & $^{11}$Four | Balanced | 64.00 | 40.17 | 60.36 | 0.234 | 0.314 |
| Moss | $^{11}$Fourier | Not Balanced | 56.80 | 51.28 | 64.4 | 0.194 | 0.443 |
| Moss | $^{11}$Fourier | Balanced | 48.0 | 39.32 | 53.6 | 0.251 | 0.408 |

*File 1 was the first half of the first night of data on each subject. Most of the training set came from file 1.

**File 2 was the second half of the first night of data on each subject. Some of the training set came from this data file.

†Correlation between entire night of sleep as machine scored against the hand scored night of sleep.

TABLE 19

CLASSIFICATION OF THE SECOND NIGHT OF RECORDED SLEEP WITH DISCRIMINANT ANALYSIS USING TRAINING DATA FROM THE FIRST NIGHT OF SLEEP

| Subject | Variables | A Priori Probabilities | 6 Categories % Correct File 3* | 6 Categories % Correct File 4** | 4 Categories % Correct Both Files | Cost Per Epoch | Correlation† |
|---|---|---|---|---|---|---|---|
| Phillips | Hist 1,4-11 & $^{11}$Fourier | Not Balanced | 43.06 | 48.57 | 50.70 | 0.202 | 0.311 |
| Phillips | Hist 1,4-11 & $^{11}$Fourier | Balanced | 41.67 | 40.0 | 47.88 | 0.224 | 0.413 |
| Phillips | $^{11}$Fourier | Not Balanced | 44.44 | 40.0 | 45.77 | 0.237 | 0.215 |
| Phillips | $^{11}$Fourier | Balanced | 33.33 | 32.68 | 38.02 | 0.315 | 0.266 |
| Moss | Hist 1,4-11 & $^{11}$Four | Not Balanced | 32.09 | 35.77 | 37.9 | 0.307 | 0.047 |
| Moss | Hist 1,4-11 & $^{11}$Four | Balanced | 26.87 | 30.08 | 33.5 | 0.367 | 0.04 |
| Moss | $^{11}$Fourier | Not Balanced | 50.0 | 47.97 | 55.65 | 0.197 | 0.325 |
| Moss | $^{11}$Fourier | Balanced | 38.81 | 40.65 | 47.58 | 0.286 | 0.239 |

*roughly the first half of the second night of data was on file 3 for each subject.

**the second half of the second night's sleep was on file 4.

†correlation between the entire night of hand scored sleep against the entire night of machine scored sleep.

TABLE 20

CLASSIFICATION OF THE FIRST NIGHT OF RECORDED

SLEEP BY A DISCRIMINANT FUNCTION

| Subject | Measures Used | A Priori Probabilities | % Correct File 1* | % Correct File 2** | % Correct Both Files | Cost Per Epochs | Correlation† |
|---|---|---|---|---|---|---|---|
| Padula | Hist & Four | Not Balanced | 59.4 | 40.65 | 54.6 | 0.195 | 0.347 |
| Padula | Hist & Four | Balanced | 50.72 | 38.2 | 50.39 | 0.261 | 0.317 |
| Padula | [11]Four | Not Balanced | 60.15 | 42.28 | 53.9 | 0.191 | 0.173 |
| Padula | [11]Four | Balanced | 44.20 | 36.83 | 39.9 | 0.334 | 0.171 |
| Gildersleeve | Hist & Four | Not Balanced | 60.50 | 57.54 | 62.5 | 0.141 | 0.554 |
| Gildersleeve | Hist & Four | Balanced | 60.50 | 50.0 | 60.27 | 0.149 | 0.519 |
| Gildersleeve | [11]Four Var | Not Balanced | 54.62 | 50.94 | 55.35 | 0.202 | 0.418 |
| Gildersleeve | [11]Four Var | Balanced | 46.21 | 27.36 | 42.86 | 0.272 | 0.358 |

*File 1 was the first half of the first night of data on each subject. Most of the training set came from file 1.

**File 2 was the second half of the first night of data on each subject. Some of the training set came from this data file.

†Correlation between entire night of sleep as machine scored against the hand scored night of sleep.

TABLE 21

CLASSIFICATION OF THE SECOND NIGHT OF RECORDED

SLEEP BY A DISCRIMINANT FUNCTION

| Subject | Measures Used | A Priori Probabilities | % Correct File 3* | % Correct File 4** | % Correct Both Files | Cost Per Epoch | Correlation† |
|---|---|---|---|---|---|---|---|
| Padula | Hist & Four | Not Balanced | 59.4 | 40.65 | 54.6 | 0.195 | 0.347 |
| Padula | Hist & Four | Balanced | 50.72 | 38.2 | 50.39 | 0.261 | 0.317 |
| Padula | [11]Four | Not Balanced | 60.15 | 42.28 | 53.9 | 0.191 | 0.173 |
| Padula | [11]Four | Balanced | 44.20 | 26.83 | 39.9 | 0.334 | 0.171 |
| Gildersleeve | Hist & Four | Not Balanced | 60.50 | 57.54 | 62.5 | 0.141 | 0.554 |
| Gildersleeve | Hist & Four | Balanced | 60.50 | 50.0 | 60.27 | 0.149 | 0.519 |
| Gildersleeve | [11]Four Var | Not Balanced | 54.62 | 50.94 | 55.35 | 0.202 | .418 |
| Gildersleeve | [11]Four Var | Balanced | 46.21 | 27.36 | 42.86 | 0.272 | 0.358 |

*roughly the first half of the second night of data was on file 3 for each subject.

**the second half of the second night's sleep was on file 4.

†correlation between the entire night of hand scored sleep against the entire night of machine scored sleep.

TABLE 22

CLASSIFICATION OF THE FIRST NIGHT OF RECORDED

SLEEP BY A DISCRIMINANT FUNCTION

| Subject | Measures Used | A Priori Probabilities | % Correct File 1* | % Correct File 2* | % Correct Both Files | Cost Per Epoch | Correlation† |
|---|---|---|---|---|---|---|---|
| Safer | Hist & Four | Not Balanced | 67.91 | 52.89 | 65.35 | 0.118 | 0.747 |
| Safer | Hist & Four | Balanced | 66.42 | 47.93 | 64.17 | 0.119 | 0.749 |
| Safer | $^{11}$Four | Not Balanced | 45.52 | 31.41 | 41.73 | 0.335 | 0.308 |
| Safer | $^{11}$Four | Balanced | 24.63 | 25.62 | 35.04 | 0.269 | 0.0461 |
| Farrington | Hist & Four | Not Balanced | 65.95 | 61.96 | 68.82 | 0.114 | 0.457 |
| Farrington | Hist & Four | Balanced | 56.38 | 53.26 | 62.37 | 0.134 | 0.468 |

*File 1 was the first half of the first night of data on each subject
Most of the training set came from file 1.

**File 2 was the second half of the first night of data on each subject.
Some of the training set came from this data file.

†Correlation between entire night of sleep as machine scored against the hand scored night of sleep.

TABLE 23

CLASSIFICATION OF THE SECOND NIGHT OF RECORDED

SLEEP WITH DISCRIMINANT ANALYSIS USING TRAINING DATA FROM THE FIRST NIGHT OF SLEEP

| Subject | Variables | A Prior Probabilities | 6 Categories % Correct File 3* | 6 Categories % Correct File 4** | 4 Categories % Correct Both Files | Cost Per Epochs | Correlations† |
|---|---|---|---|---|---|---|---|
| Safer | Hist & Four | Not Balanced | 33.33 | 44.72 | 39.6 | 0.402 | 0.381 |
| Safer | Hist & Four | Balanced | 32.61 | 41.46 | 37.69 | 0.407 | 0.375 |
| Safer | [11]Four | Not Balanced | 27.53 | 28.46 | 30.36 | 0.503 | 0.122 |
| Safer | [11]Four | Balanced | 5.79 | 20.32 | 20.0 | 0.536 | 0.062 |
| Farrington | Hist & Four | Not Balanced | 59.38 | 55.91 | 65.08 | 0.133 | 0.639 |
| Farrington | Hist & Four | Balanced | 52.08 | 46.23 | 58.2 | 0.149 | 0.632 |

*roughly the first half of the second night of data was on file 3 for each subject.

**the second half of the second night's sleep was on file 4

†correlation between the entire night of hand scored . nst the entire night of machine scored sleep.

TABLE 24

PROBABILITY OF OCCURRENCE, P

| | A | 1 | 2 | 3 | 4 | REM |
|---|---|---|---|---|---|---|
| Safer | .237 | .072 | .338 | .072 | .072 | .208 |
| Schmidt | .090 | .090 | .423 | .090 | .090 | .208 |
| Farrington | .087 | .087 | .434 | .087 | .087 | .217 |
| Chinoy | .088 | .088 | .377 | .079 | .088 | .281 |
| Gildersleeve | .082 | .082 | .418 | .082 | .098 | .237 |
| Moss | .095 | .103 | .465 | .086 | .000 | 250 |
| Verick | .076 | .076 | .481 | .076 | .091 | .[illegible]8 |
| Nordyke | .088 | .088 | .368 | .088 | .158 | .210 |
| Phillips | .109 | .099 | .446 | .099 | .099 | .148 |
| Padula | .075 | .090 | .466 | .075 | .075 | .218 |

$$P_I = \frac{\text{Number of Epochs of Stage I in Training Set}}{\text{Total Number of Training Epochs}}$$

I = Awake, 1, 2, 3, 4, and REM

Tables 25, 26, 27 and 28 represent examples of good results using the discriminant function procedure.

Tables 25 and 26 represent results obtained from an attempt to classify the first and second nights of subject Chinoy's sleep with variables 1,2,3,5,7, 9 and 11. A graphical representation of his sleep pattern for the two nights based on EEG hand-scored records is presented in Figures 3 and 4. The corresponding heart rate machine scored sleep patterns are shown in Figures 5 and 6.

Tables 27 and 28 give the details of an analysis of the same data using a combined histogram Fourier analysis measure set. The computer-generated classification based on these measures is presented in Figure 7 (first night of recorded sleep) and Figure 8 (second night of recorded sleep).

TABLE 25

EXAMPLE OF CLASSIFICATION RESULTS FOR FIRST NIGHT OF RECORDED SLEEP

Subject: Chinoy
Variables: 1,2,3,5,7,9,11
Night: 3 (first recorded night of sleep)
A priori probability: Actual frequency of occurrence

This stage of sleep..........

| was classified as | Awake | 1 | 2 | 3 | 4 | REM |
|---|---|---|---|---|---|---|
| Awake | 4 | 0 | 2 | 0 | 0 | 2 |
| 1 | 1 | 8 | 2 | 1 | 0 | 4 |
| 2 | 0 | 1 | 66 | 3 | 12 | 6 |
| 3 | 0 | 0 | 6 | 1 | 0 | 0 |
| 4 | 1 | 0 | 2 | 0 | 12 | 0 |
| REM | 1 | 4 | 5 | 0 | 0 | 56 |

REDUCED SLEEP MATRIX

This stage of sleep............

| was classified as | Awake | (1,REM) | 2 | (3,4) |
|---|---|---|---|---|
| Awake | 4 | 2 | 2 | 0 |
| (1,REM) | 2 | 72 | 7 | 1 |
| 2 | 0 | 7 | 66 | 15 |
| (3,4) | 1 | 0 | 8 | 13 |

TABLE 26

EXAMPLE OF CLASSIFICATION RESULTS FOR SECOND NIGHT OF RECORDED SLEEP

| | |
|---|---|
| Subject: | Chinoy |
| Variables: | 1,2,3,5,7,9,11 |
| Night: | 4 (second recorded night of sleep) |
| A priori probability: | Actual frequency of occurrence |

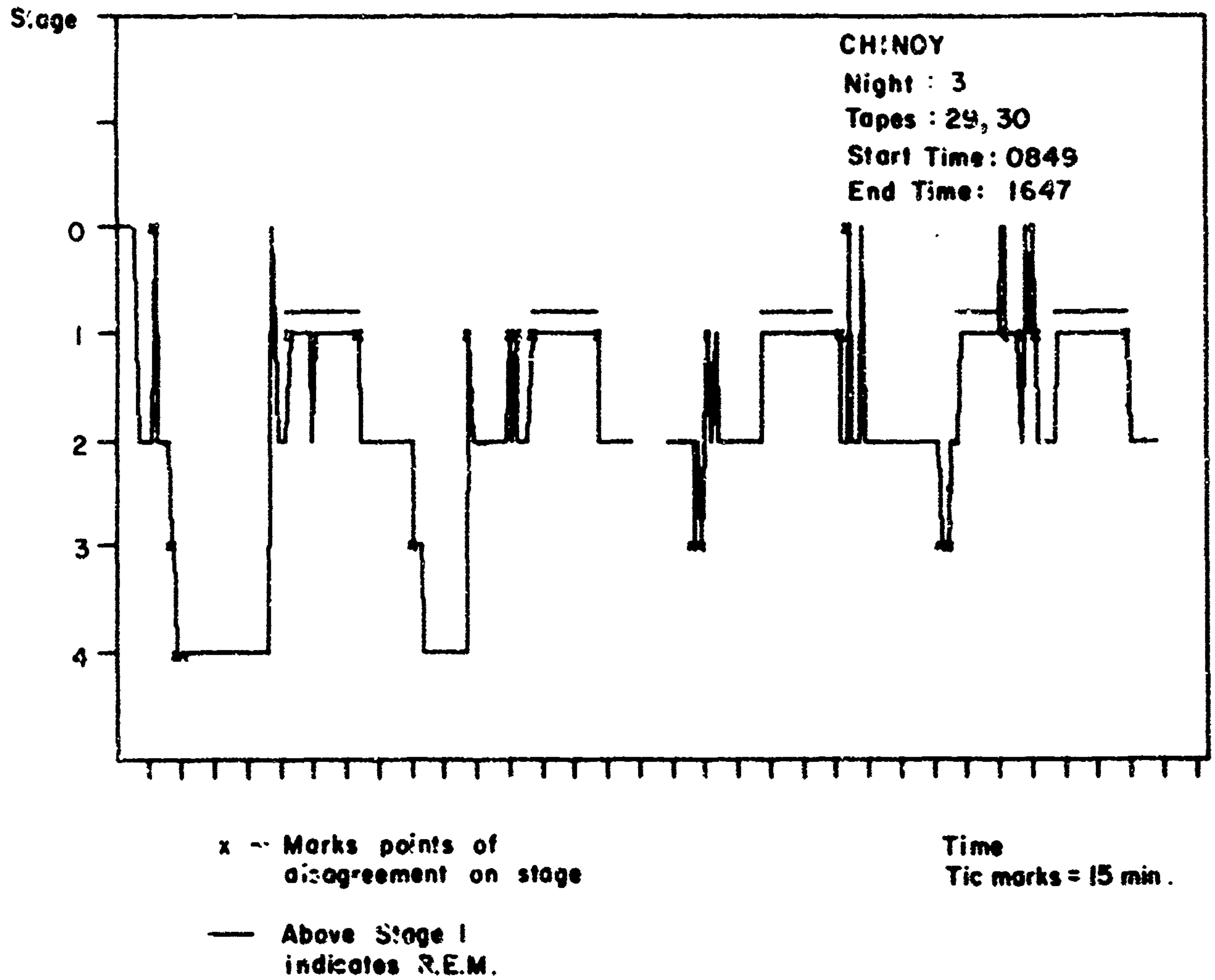
This stage of sleep..............

| . . . was classified as | | Awake | 1 | 2 | 3 | 4 | REM |
|---|---|---|---|---|---|---|---|
| | Awake | 5 | 0 | 1 | 0 | 0 | 5 |
| | 1 | 1 | 1 | 3 | 0 | 0 | 10 |
| | 2 | 0 | 2 | 78 | 3 | 14 | 16 |
| | 3 | 1 | 1 | 4 | 1 | 2 | 3 |
| | 4 | 0 | 0 | 7 | 0 | 0 | 0 |
| | REM | 0 | 1 | 1 | 0 | 0 | 27 |

REDUCED SLEEP MATRIX

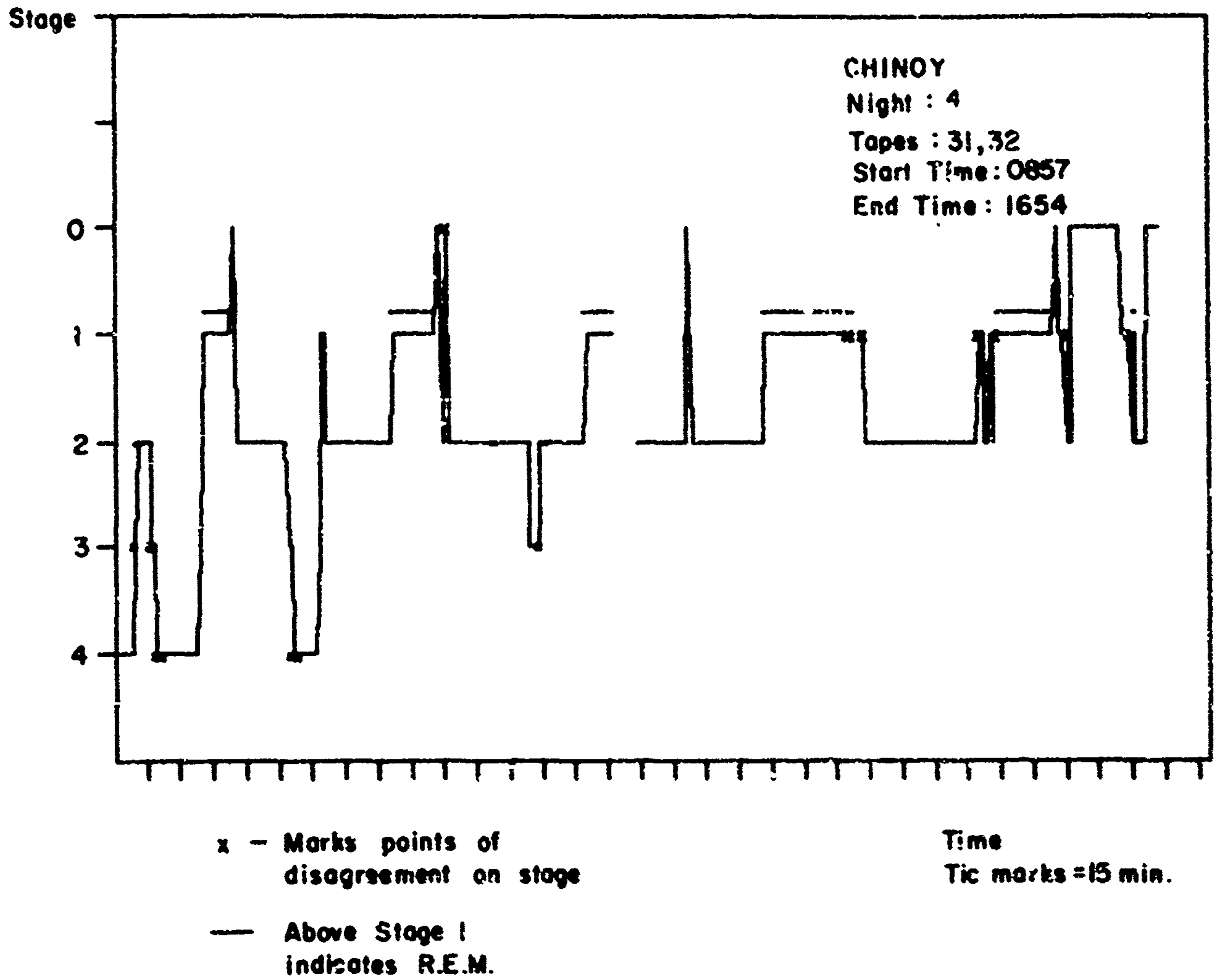
This stage of sleep.............

| . . . was classified as | | Awake | (1,REM) | 2 | (3,4) |
|---|---|---|---|---|---|
| | Awake | 5 | 5 | 1 | 0 |
| | (1,REM) | 1 | 39 | 4 | 0 |
| | 2 | 0 | 18 | 78 | 17 |
| | (3,4) | 1 | 4 | 11 | 3 |



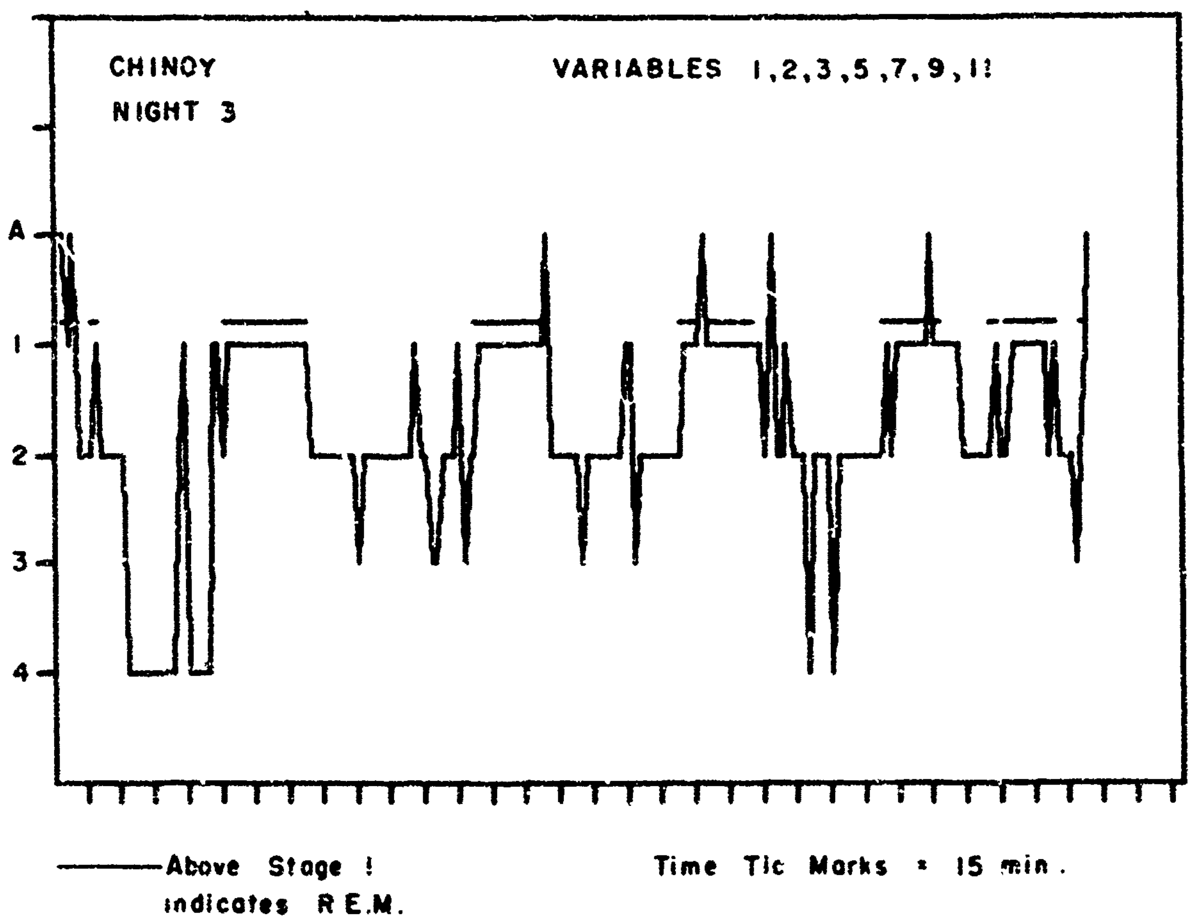

EEG
SLEEP PATTERNS

Figure 3



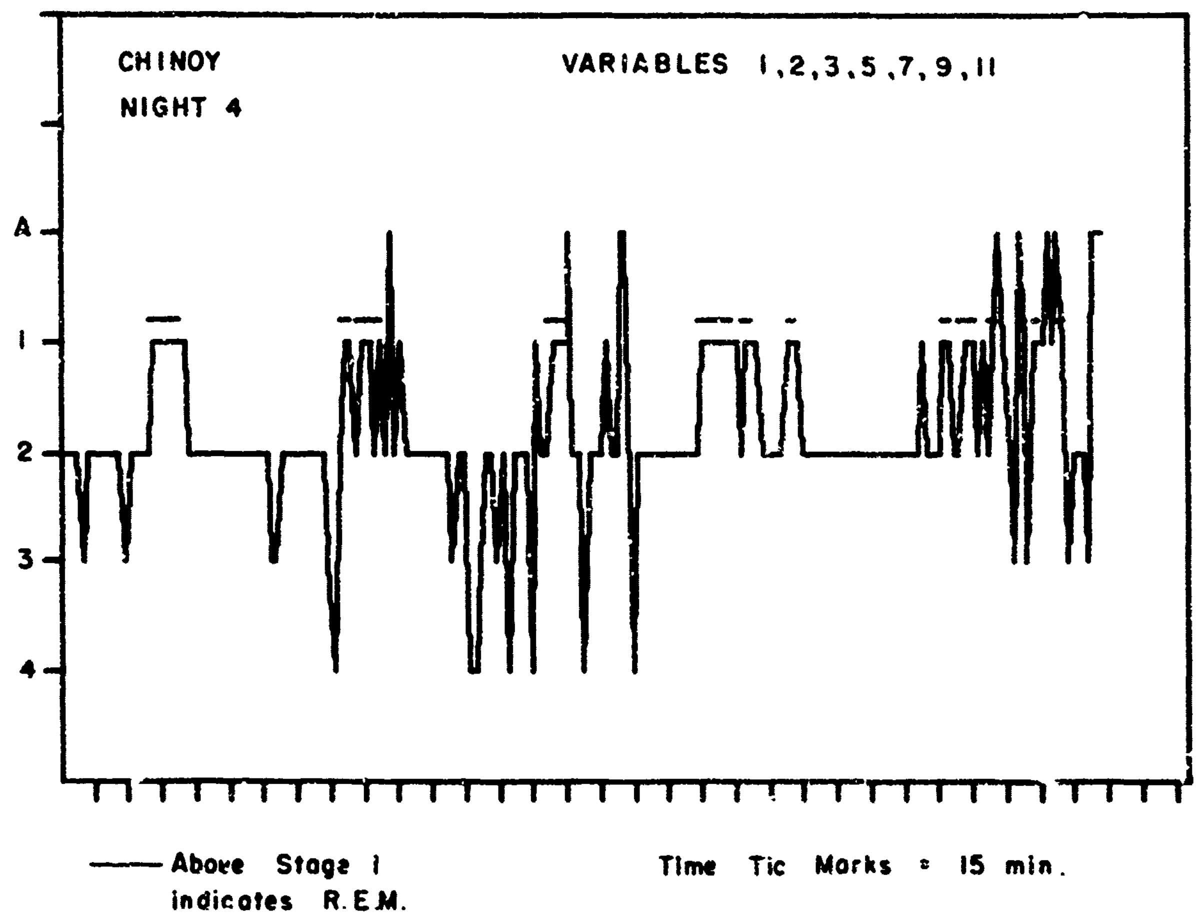

EEG
SLEEP PATTERNS

Figure 4




**CLASSIFICATION OF HEART RATE DATA INTO SLEEP STAGES**

Figure 5




CLASSIFICATION OF HEART RATE DATA INTO SLEEP STAGES

Figure 6

## TABLE 27

## EXAMPLE OF CLASSIFICATION RESULTS FOR FIRST NIGHT OF RECORDED SLEEP

Subject: Chinoy
Variables: 1, 4-11 and eleven Fourier
Night: 3
A priori probability: Actual frequency of occurrence

This stage of sleep............

| . . . was classified as | | Awake | 1 | 2 | 3 | 4 | REM |
|---|---|---|---|---|---|---|---|
| | Awake | 4 | 2 | 2 | 1 | 0 | 1 |
| | 1 | 0 | 3 | 5 | 0 | 0 | 3 |
| | 2 | 0 | 4 | 55 | 3 | 14 | 6 |
| | 3 | 0 | 0 | 3 | 1 | 1 | 0 |
| | 4 | 1 | 0 | 5 | 0 | 9 | 0 |
| | REM | 2 | 4 | 5 | 0 | 0 | 58 |

### REDUCED SLEEP MATRIX

This stage of sleep.............

| . . . was classified as | | Awake | (1, REM) | 2 | (3,4) |
|---|---|---|---|---|---|
| | Awake | 4 | 3 | 2 | 1 |
| | (1,REM) | 2 | 68 | 10 | 0 |
| | 2 | 0 | 10 | 55 | 17 |
| | (3,4) | 1 | 0 | 8 | 11 |

TABLE 28

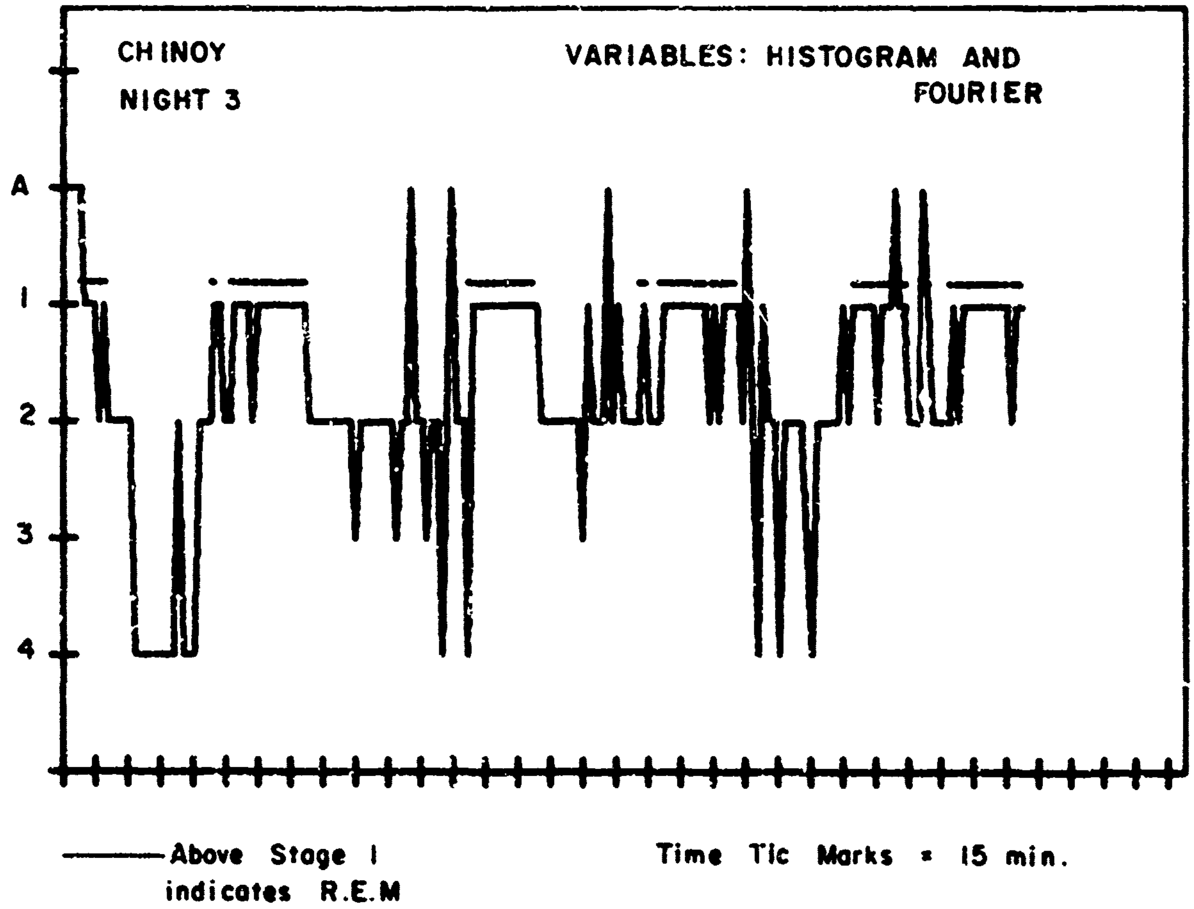
EXAMPLE OF CLASSIFICATION RESULTS FOR SECOND NIGHT OF RECORDED SLEEP

Subject: Chinoy
Variables: 1, 4-11 and eleven Fourier
Night: 4
A priori probability: Actual frequency of occurrence

| | | This stage of sleep............... | | | | | |
|---|---|---|---|---|---|---|---|
| | | Awake | 1 | 2 | 3 | 4 | REM |
| was classified as | Awake | 0 | 0 | 1 | 1 | 0 | 3 |
| | 1 | 1 | 1 | 9 | 0 | 1 | 11 |
| | 2 | 1 | 4 | 74 | 3 | 12 | 17 |
| | 3 | 0 | 0 | 2 | 0 | 2 | 0 |
| | 4 | 0 | 0 | 6 | 0 | 1 | 0 |
| | REM | 0 | 0 | 1 | 0 | 0 | 30 |

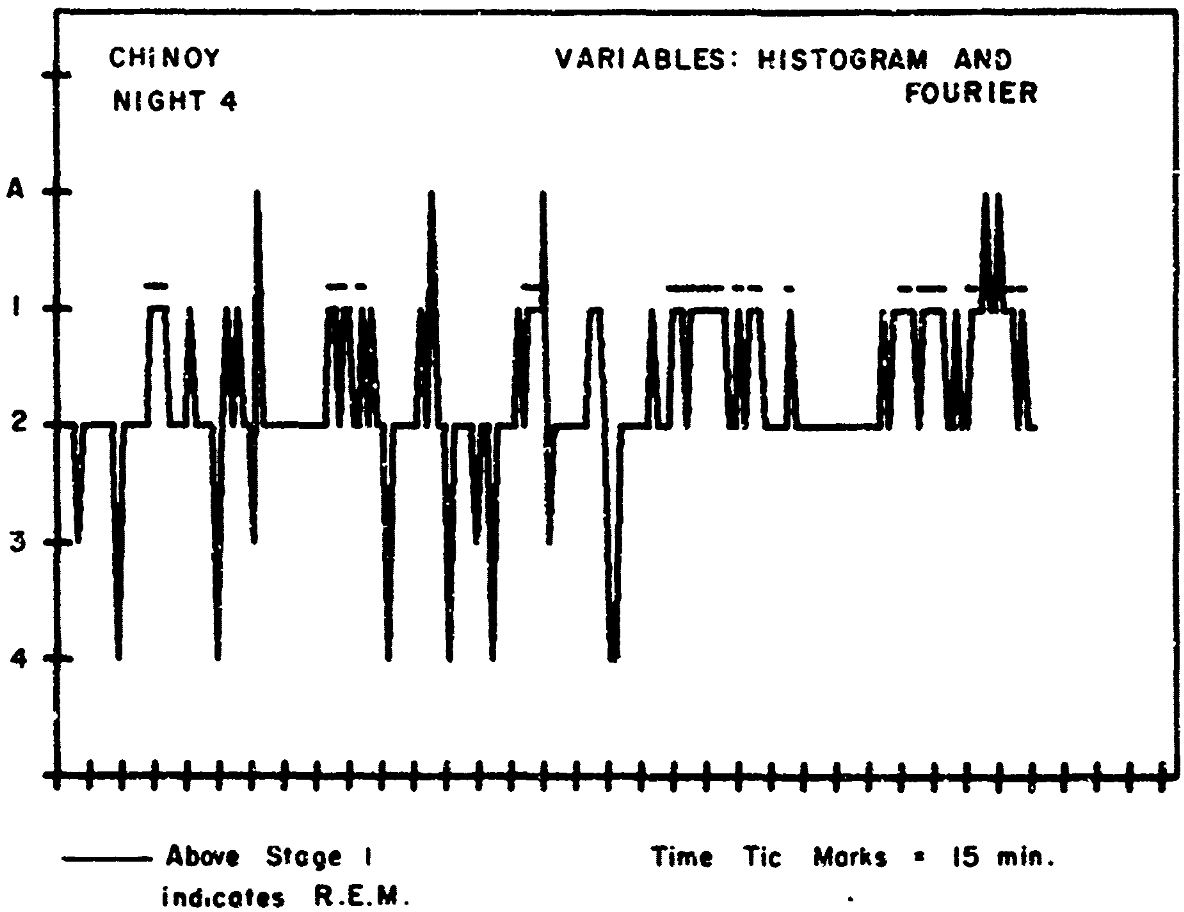
REDUCED SLEEP MATRIX

| | | This stage of sleep................ | | | |
|---|---|---|---|---|---|
| | | Awake | (1,REM) | 2 | (3,4) |
| was classified as | Awake | 0 | 3 | 1 | 1 |
| | (1,REM) | 1 | 42 | 10 | 1 |
| | 2 | 1 | 21 | 74 | 15 |
| | (3,4) | 0 | 0 | 8 | 3 |




CLASSIFICATION OF HEART RATE DATA INTO SLEEP STAGES

Figure 7




CLASSIFICATION OF HEART RATE DATA
INTO SLEEP STAGES

Figure 8

DISCUSSION

The overall results of a classification by discriminant analysis were somewhat discouraging when the second night of data was considered. As might be expected the unbalanced a priori probabilities consistently produced better classification results.

Much of the inability of this algorithm to correctly classify sleep is rooted in the fact that there was considerable amount of intra-subject variation and even intra-night variation in the beat-to-beat heart rate. Aldredge, et al (3) has investigated the intra-subject and intra-cycle variation in the mean heart rate and sample standard deviation of the data used for this investigation. They examined the possibility that the mean values of average heart rate and sample standard deviation might not be consistent throughout a night of sleep or between two nights of sleep. Undesired variations in the average and sample standard deviation in heart rate during a night of sleep were determined by testing the following hypothesis with analysis of variances for each subject:

$H_o$: The mean values of a random variable X for each stage I are equal for all cycles of sleep during a single night of sleep.

(X is either equal average heart rate or sample standard deviation, and I represents either (1,REM), 2, or (3,4)).

Aldredge, et al, concluded that for a single night of sleep the above hypothesis could be rejected in most of the cases at .0001 significance level when X represented the average heart rate and I was either equal to (1,REM) or 2 or (3,4). However, the hypothesis could not be rejected at the .0001 level for any stage of sleep when the sample standard deviation was tested. A close inspection of the intr-cycle variation in average heart rate suggested that the heart rate during a cycle of sleep was influenced by the average REM heart rate at the onset of the cycle. The null hypothesis of equal means for average heart rate values of stage REM and of stage 2 for any given cycle was rejected in favor of the alternate hypothesis that the mean for REM was greater than the mean for stage 2 at the .05 significance level. Also the alternate hypothesis was accepted when the mean values of mean averaged heart rate values of stage REM are greater than that of combined stage REM and combined stages 3 and 4 were compared with a one tailed t-test at a 5% significance level. A third null hypothesis which stated that the mean averaged heart rate values of stage 2 were equal to that of combined stages 3 and 4 could not be rejected.

The performance of the Fourier measure and their ability to represent sleep information was disappointing. With the exception of Messrs, Paduia and Moss, the Fourier measures alone did not improve the cost of sleep scoring. When combined with the histogram measures, five of the subjects had a slight decrease in cost of classification. Table 29 illustrates the relative cost of of using histogram measures, Fourier measures

TABLE 29

COST FOR VARIOUS MEASURES
(Second Night)

| Subject | Histogram Measures | Fourier Measures | Combined Fourier and Histogram Measures |
|---|---|---|---|
| Chinoy | .12 | .25 | 11 |
| Safer | .39 | .50 | .40 |
| Gildersleeve | .15 | .20 | .17 |
| Padula | .50 | .27 | .48 |
| Moss | .35 | .19 | .30 |
| Phillips | .21 | .23 | .20 |
| Schmidt | .25 | .31 | .27 |
| Farrington | .14 | .25 | .13 |

alone and the combined histogram and Fourier measures.

Examination of the summary tables indicated that all nine histograms were never used in the same covaraince matrix since the linear dependence of one histogram variable on the other histogram variables produces a singular matrix which is non-invertible.

Table 30 considers that advantages of using balanced and unbalanced a priori probabilities. In terms of our empirically-derived cost function, the non-balanced a priori probabilities are quite similar. If percent of accuracy diagnosed sleep epochs is the criterion, the non-balanced a priori probabilities are decidedly superior.

CONCLUSIONS

Result of this research indicates that it is possible to classify heart rate patterns into sleep stages. However, the results are not overwhelming, in spite of the fact that the analysis of variance indicates an optimistic possibility of sleep stage classification ability of these measures. Much of the difficulty experienced by this algorithm can be attributed to intra-night, intra-subject variations in mean value as the 90-minutes sleep cycles progress throughout the night.

Our empirically derived cost function proved to be a useful measure of the effectiveness of our sleep classification algorithm. We contend that any measure of merit used to evaluate a device or procedure should of necessity be closely related to the original problem, in this case the study of sleeping patterns.

TABLE 30

EFFECT OF USING A PRIORI PROBABILITIES
IN CLASSIFICATION ALGORITHM

| Classification Criteria | Number of Classification Runs | |
|---|---|---|
| I Lowest Cost Using | Night 1 | Night 2 |
| *A Priori* Probability | 14 | 14 |
| Balanced Probabilities (likelihood ratio) | 0 | 2 |
| No Significant Difference Between Above Methods | 4 | 4 |
| II Highest Correlation Using | | |
| *A Priori* Probability | 8 | 4 |
| Likelihood Ratio | 7 | 7 |
| No Difference | 4 | 8 |
| III Percent Accuracy Using | | |
| *A Priori* Probability | 15 | 13 |
| Likelihood Ratio | 0 | 0 |
| No Difference | 3 | 5 |

We also conclude that if it were possible to normalized the heart rate data so that the inter (90 minutes) sleep cycle variation in mean were less, then better results would have been obtained. We suggest that if an alternate algorithm could be developed to determine the beginning of each 90 minute sleep cycle, then this algorithm would be able to accurately classify the data into sleep stages.

Also, if it were possible to extract respiration information from amplitude variations in the QRS complex, this additional variable might improve the accuracy of the algorithm.

REFERENCES


1. Agnew, Jr., H. W., Webb, W. B., Williams, R. L., The First Night Effect: An EEG Study of Sleep, Psychophysiology, 1966, 2:263-266.

2. Agnew, H. W., Personal Communication.

3. Aldredge, Joanne L., A. J. Welch, Philip C. Richardson, F. B. Vogt, The Extraction of Sleep Information From Heart Rate Data: Analysis of the Sleep Cycle, TR. 107, Electronics Research Center, Austin, Texas 1971.

4. Berry, C. A.: Preliminary Clinical Report of the Medical Aspects of Apollos VII and VIII. Aerospace Medicine 40:245-254, 1969.

5. Baust, Walter and Bernd Bohnert. The Regulation of Heart Rate During Sleep. Exp. Brain Research, 7:169-180, 1969.

6. Brooks, C. McC. et al. Sleep and Variations in Certain Functional Activities Accompanying Cyclic Changes in Depth of Sleep. J. Appl. Physiol., 9:97-107, 1956.

7. Cooley, W. W., and P. R. Lohnes: Multivariate Procedures for the Behavioral Sciences. John Wiley and Sons, Inc., 1962.

8. Dement, W., and N. Kleitman: Cyclic Variations in EEG During Sleep and Their Relation to Eye Movements, Body Motility, and Dreaming. EEG Clin. Neurophysiol. 9:673-690, 1957.

9. Fisher, C., and W. Dement: Studies on the Psychopathology of Sleep and Dreams. Am. J. Psychiat. 119:1160-1168, 1963.

10. Hartman, B. O., and G. K. Cantrell: Mol: Crew Performance on Demanding Work/Rest Schedules Compounded by Sleep Deprivation. SAM-TR-67-99: 1-29, 1967.

11. Khatri, Ibrahim M. and Edward D. Freis. Hemodynamic Changes During Sleep. J. Appl. Physiol., 22(5): 867-873, 1967.

12. Kleitman, N.: Sleep. Sci. Am. 187:34-38, 1952.

13. Kollar, E. J., R. O. Pasnau, R. T. Rubin, P. Naitoh, G. G. Slater and A. Kales: Psychological, Psychophysiological and Biochemical Correlates of Prolonged Sleep Deprivation. Amer. J. Psychiat. 126:70-79, 1969.

14. Lessard, Charles S., Ford, G. E., Hughes, H. M., Decomposition and Variability of EEG Information During Sleep, Twelfth Midwest Symposium on Circuit Theory, 1969, 7.

15. MacWilliams, J. A., Blood Pressure and Heart Action in Sleep and Dreams. Brit. Med. J. 2:1196-1200, 1923.

16. Monroe, L. J.: Inter-rate Reliability and the Role of Experience in Scoring EEG Sleep Record: Phase 1. Psychophysiology 5:376-384, 1969.

17. Rao, C. R.: Advanced Statistical Methods in Biometric Research. Wiley, New York, 286-328, 1952.

18. Rodin, E. A., Luby, E. D. and Gottlieb, J. S., The Electroencephalogram During Prolonged Experimental Sleep Deprivation. Electroenceph. Clin. Neurophysiol., 1962, 14:544-551.

19. Simons, David G. and Neal E. Lof. Computed Profiles of Respiratory Rate Values Observed Through a Spectrum of Aerospace Stresses. U.S.A.F. School of Aerospace Medicine, Brooks A.F.B., Texas.

20. Snyder, Frederick, et al. Changes in Respiration, Heart Rate, and Systolic Blood Pressure in Human Sleep. J. Appl. Physiol., 19:417-422, 1964.

21. Welch, A. J., P. C. Richardson, C. W. Thomas and J. M. Aldredge. Bandwidth Reduction of Sleep Information. Technical Report No. 92, Electronics Research Center, University of Texas, 1970.

22. West, L. J., H. Janszen, B. K. Lester and F. S. Cornelisoon, Jr.: The Psychosis of Sleep Deprivation. Ann. N. Y. Acad. Sci. 96:66-70, 1962.

23. Williams, R. L., H. W. Agnew, Jr., and W. B. Webb: Effects of Prolonged Stage Four and 1-Rem Sleep Deprivation - EEG, Task Performance, and Psychologic Responses. SAM-TR-67-59: 1-10, July 1967.

24. Williams, R. L., H. W. Agnew, Jr., and W. B. Webb: Sleep Patterns in Young Adults: An EEG Study. Electroenceph. Clin. Neurophysiol. 17:376-381, 1964.

25. Veldman, D. J.: Fortran Programming for the Behavioral Sciences. Holt, Rinehart, and Winston, New York, 1967.

## Modif'ed Discriminant Analysis

The original discriminant analysis program is that of Dr. Donald J. Veldman of The University of Texas at Austin, Department of Educational Psychology and is well documented in his book Fortran Programming for the Behavioral Sciences. In addition to his discriminant analysis program, we have found it helpful to use some of the minor subroutines included in his work; namely PRTS and PCDS, which are efficient print and punch subroutines and CORS which computes means, sigmas, and intercorrelations. The discriminant program determines a transformation matrix of data from known groups (sleep stages) which maximizes the distance between centroid while holding the overall distance between points of data constant. The dimensionality of the new space is the minimum of (1) the number of groups minus one, or (2) the number of variables. Basically discriminant analysis is a transformation from the data space to a new reduced measure space.

The classification program of Cooley and Lohnes requires not only the D-WTS matrix and centroids, but the covariance matrix at each group. The modifications of the discriminant analysis program for computing the covariance matrices are described in the following paragraphs.

In loop 35 of the program original SDSCRIM variable sums and cross products are accumulated for all samples of the variables (i.e. all data pooled regardless of group classification). After storing the sum of the score for each group in S and the sum of the cross products in A, a "within group matrix" is computed and added to W. It is this "within group matrix" for each group which we desire to use to calculate the dispersion matrix (i.e. covariance matrix) for each group. We wish to save that matrix for each group before it is added with those of the other groups into matrix W. This is done in the modified version called JDSCRIM in line c by saving matrix B in CC which will be added in to W in statement 34. The triangular matrix we wish is now available in B(I.J.); we need to divide it again by C(M) [statement d] and fill in the portion below the diagonal [statement e]. If we wish to calculate and punch the dispersion matrix for each group a minus one is placed in cc 24-25 of the control card. The matrix B for each group is written onto scratch tape 1 in loop 33. After the D-WTS are calculated (through statement 60) we have the items necessary for calculating the reduced space dispersion matrices which will be done in the subroutine DMG, called after statement 60.

```
      PROGRAM JDISC(INPUT,OUTPUT,PUNCH,TAPE1,TAPE2)
      CALL JDSCRIM
      CALL EXIT  $  END
      SUBROUTINE JDSCRIM
     0DIMENSION A(70,70), W(70,70), C(70,70), S(70,25), T(70),
     1 V(70), X(70), Y(70), Z(70), Q(70), G(25), KF(16), KH(15)
      DIMENSION  B(70,70),CC(70,70)
      N1 = 70  $ N2 = 25
    5 CALL CCDS (KF, NV, NG, KW, KT, KEY)
      IF(KEY.EQ.-1) REWIND 1
      CALL INPUT (ID, X, 0, KF, NV)
      DO 10 I = 1,NV
      DO 10 J = I,NV
      C(I,J) = 0.0
   10 W(I,J) = 0.0
      IF (KT .GT. 0) REWIND 2
      DO 35 M = 1,NG
      READ 15, N, KH
   15 FORMAT (I5, 15A5)
      PRINT 20, M, N, KH
   20 FORMAT (/ 6H GROUP, I2, I8, 10H SUBJECTS., 2X, 15A5)
      G(M) = N
      DO 25 I = 1,NV
      S(I,M) = 0.0
      DO 25 J = I,NV
   25 A(I,J) = 0.0
      DO 30 I = 1,N
      CALL INPUT (ID, X, N + M * 1000, KF, NV)
      IF (KT .GT. 0) WRITE (2) ID, (X(J), J = 1,NV)
      DO 30 J = 1,NV
      S(J,M) = S(J,M) + X(J)
      DO 30 K = J,NV
   30 A(J,K) = A(J,K) + X(J) * X(K)
      DO 34 I=1,NV
      DO 34 J=I,NV
      C(I,J) = C(I,J) + A(I,J)
      B(I,J)=              (A(I,J) - S(I,M) * S(J,M) /G(M))
      CC(I,J)= B(I,J)
      B(I,J) = B(I,J)/G(M)
      B(J,I) =B(I,J)
   34 W(I,J) = W(I,J) + CC(I,J)
      IF (KEY.NE.-1) GO TO 35
      DO 33 I=1,NV
   33 WRITE (1) (B(I,J),J=1,NV)
   35 CONTINUE
      TN = SUMF(G, 1, NG, N2)
      DO 40 I = 1,NV
      T(I) = SUMF(S, -I, NG, N1) / TN
   40 Q(I) = C(I,I)
      DO 45 I = 1,NV
      DO 45 J = I,NV
      C(I,J) = C(I,J) / TN - T(I) * T(J)
      C(J,I) = C(I,J)
      A(I,J) = C(I,J) * TN - W(I,J)
      A(J,I) = A(I,J)
   45 W(J,I) = W(I,J)
      CALL INVS (NV, W, X, Y, Z, N1)
      DO 55 I = 1,NV
      DO 50 J = 1,NV
```

```
50 X(J) = W(I,J)
   DO 55 J = 1,NV
55 W(I,J) = SCPF(X, A, 1, J, NV, N1)
   NF = MINO(NG - 1, NV)
   CALL AEVS (NV, NF, 0.0, W, A, V, X, Y, Z, N1)
   DO 60 J = 1,NF
   E = 1.0 / SQRT(V(J))
   DO 60 I = 1,NV
60 A(I,J) = A(I,J) * E
   IF (KW .EQ. 1) CALL PCDS (A, NV, NF, 5HD WTS, N1)
   IF (KW .EQ. 1) CALL PRTS (A, NV, NF, 5HD WTS, N1)
   IF(KEY.EQ.-1)CALL DMG(A,NV,NF,NG,N1)
   DO 65 I = 1,NV
65 X(I) = SQRT(C(I,I))
   CALL AXBS (C, A, W, NV, NF, NV, N1)
   DO 70 I = 1,NF
70 Y(I) = SQRT(SCPF(A, W, I, I, NV, N1))
   DO 75 I = 1,NV
   DO 75 J = 1,NF
75 C(I,J) = W(I,J) / (X(I) * Y(J))
   TR = SUMF(V, 1, NF, N1)
   XL = 1.0
   DO 80 I = 1,NF
   X(I) = V(I) / TR * 100.0
80 XL = XL * (1.0 / (1.0 + V(I)))   $VN=NV$GN=NG$GM=GN-1.0
   SS = SQRT((VN**2 * GM**2 - 4.0) / (VN**2 + GM**2 - 5.0))
   YY = XL**(1.0 / SS)
   FA = VN * GM
   FB = ((TN - 1.0) - (VN + GN) / 2.0) * SS - (VN * GM - 2.0) / 2.0
   F = (FB * (1.0 - YY)) / (YY * FA)
   P = PRBF(FA, FB, F)
   PRINT 85, XL, FA, FB, F, P
850FORMAT (// 15H WILKS LAMBDA =, F10.3 // 7H D.F. =, F5.0,
  14H AND, F7.0 // 10H F-RATIO =, F8.3, 5X, 3HP =, F7.4)
   DF = VN + GN
   CC = TN - DF / 2.0
   DO 90 I = 1,NF
   CS = CC * ALOG(1.0 + V(I))
   DF = DF - 2.0
   P = PRBF(DF, 1000.0, CS / DF)
90 PRINT 95, I, X(I), CS, DF, P
950FORMAT (/ 5HOROOT, I2, F10.2, 14H PCT. VARIANCE //
  113H CHI-SQUARE =, F10.3, 5X, 6HD.F. =, F5.0, 5X, 3HP =, F7.4)
   DO 100 I = 1,NV
   T(I) = T(I) * TN
   DO 100 J = 1,NG
100 S(I,J) = S(I,J) / G(J)
   CALL AXBS (S, A, W, -NC, NF, NV, N1)
   CALL PRTS (W, NG, NF, 5HCENT., N1)
   CALL PCDS (W, NG, NF, 5HCENT., N1)
   CALL PRTS (C, NV, NF, 6HCOREL., N1)
   DFW = TN - GN
   PRINT 105, GM, DFW
1050FORMAT (// 26H UNIVARIATE F-TESTS, DFB =, F3.0,
  15H DFW =, F6.0 /// 18H VARIABLE  F-RATIO, 6X, 1HP)
   DO 115 I = 1,NV
   B = 0.0
   DO 110 J = 1,NG
110 B = B + S(I,J)**2 * G(J)
   CC = T(I)**2 / TN
```

```
      F = ((B - CC) * DFW) / ((Q(I) - B) * GM)
      P = PRBF(GM, DFW, F)
  115 PRINT 120, I, F, P
  120 FORMAT (/ I6, F12.4, F10.4)
      CALL PRTS (S, NV, NG, 6HG MEAN, N1)
      IF (KT .EQ. 0) GO TO 5
      REWIND 2     $ NT = TN
      DO 130 I = 1,NT
      READ (2) ID, (X(J), J = 1,NV)
      DO 125 J = 1,NF
  125 Y(J) = SCPF(X, A, 1, J, NV, N1)
  130 CALL SUBS (Y, NF, 2HCS, ID)
      GO TO 5   $  RETURN
      END
```

```
      SUBROUTINE SDISCRIM
     0DIMENSION A(70,70), W(70,70), C(70,70), S(70,25), T(70),
     1 V(70), X(70), Y(70), Z(70), Q(70), G(25), KF(16), KH(15)
      DIMENSION  B(70,70),CC(70,70)
      N1 = 70  $ N2 = 25
    5 CALL CCDS (KF, NV, NG, KW, KT, KEY)
      CALL INPUT (ID, X, 0, KF, NV)
      DO 10 I = 1,NV
      DO 10 J = 1,NV
      C(I,J) = 0.0
   10 W(I,J) = 0.0
      IF (KT .GT. 0) REWIND 2
      DO 35 M = 1,NG
      READ 15, N, KH
   15 FORMAT (I5, 15A5)
      PRINT 20, M, N, KH
   20 FORMAT (/ 6H GROUP, I2, I8, 10H SUBJECTS., 2X, 15A5)
      G(M) = N
      DO 25 I = 1,NV
      S(I,M) = 0.0
      DO 25 J = 1,NV
   25 A(I,J) = 0.0
      DO 30 I = 1,N
      CALL INPUT (ID, X, N + M * 1000, KF, NV)
      IF (KT .GT. 0) WRITE (2) ID, (X(J), J = 1,NV)
      DO 30 J = 1,NV
      S(J,M) = S(J,M) + X(J)
      DO 30 K = J,NV
   30 A(J,K) = A(J,K) + X(J) * X(K)
      DO 35 I = 1,NV
      DO 35 J = I,NV
      C(I,J) = C(I,J) + A(I,J)
   35 W(I,J) = W(I,J) + (A(I,J) - S(I,M) * S(J,M) / G(M))
      TN = SUMF(G, 1, NG, N2)
      DO 40 I = 1,NV
      T(I) = SUMF(S, -I, NG, N1) / TN
   40 Q(I) = C(I,I)
      DO 45 I = 1,NV
      DO 45 J = I,NV
      C(I,J) = C(I,J) / TN - T(I) * T(J)
      C(J,I) = C(I,J)
      A(I,J) = C(I,J) * TN - W(I,J)
      A(J,I) = A(I,J)
   45 W(J,I) = W(I,J)
      CALL INVS (NV, W, X, Y, Z, N1)
      DO 55 I = 1,NV
      DO 50 J = 1,NV
   50 X(J) = W(I,J)
      DO 55 J = 1,NV
   55 W(I,J) = SCPF(X, A, 1, J, NV, N1)
      NF = MIN0(NG - 1, NV)
      CALL AEVS (NV, NF, 0.0, W, A, V, X, Y, Z, N1)
      DO 60 J = 1,NF
      E = 1.0 / SQRT(V(J))
      DO 60 I = 1,NV
   60 A(I,J) = A(I,J) * E
      IF (KW .EQ. 1) CALL PCDS (A, NV, NF, 5HD WTS, N1)
      IF (KW .EQ. 1) CALL PRTS (A, NV, NF, 5HD WTS, N1)
      DO 65 I = 1,NV
```

```
65 X(I) = SQRT(C(I,I))
   CALL AXBS (C, A, W, NV, NF, NV, N1)
   DO 70 I = 1,NF
70 Y(I) = SQRT(SCPF(A, W, I, I, NV, N1))
   DO 75 I = 1,NV
   DO 75 J = 1,NF
75 C(I,J) = W(I,J) / (X(I) * Y(J))
   TR = SUMF(V, 1, NF, N1)
   XL = 1.0
   DO 80 I = 1,NF
   X(I) = V(I) / TR * 100.0
80 XL = XL * (1.0 / (1.0 + V(I)))    $VN=NV$GN=NG$GM=GN-1.0
   SS = SQRT((VN**2 * GM**2 - 4.0) / (VN**2 + GM**2 - 5.0))
   YY = XL**(1.0 / SS)
   FA = VN * GM
   FB = ((TN - 1.0) - (VN + GN) / 2.0) * SS - (VN * GM - 2.0) / 2.0
   F = (FB * (1.0 - YY)) / (YY * FA)
   P = PRBF(FA, FB, F)
   PRINT 85, XL, FA, FB, F, P
85 FORMAT (// 15H WILKS LAMBDA =, F10.3 // 7H D.F. =, F5.0,
  14H AND, F7.0 // 10H F-RATIO =, F8.3, 5X, 3HP =, F7.4)
   DF = VN + GN
   CC = TN - DF / 2.0
   DO 90 I = 1,NF
   CS = CC * ALOG(1.0 + V(I))
   DF = DF - 2.0
   P = PRBF(DF, 1000.0, CS / DF)
90 PRINT 95, I, X(I), CS, DF, P
95 FORMAT (/ 5HROOT, I2, F10.2, 14H PCT. VARIANCE //
  113H CHI-SQUARE =, F10.3, 5X, 6HD.F. =, F5.0, 5X, 3HP =, F7.4)
   DO 100 I = 1,NV
   T(I) = T(I) / TN
   DO 100 J = 1,NG
100 S(I,J) = S(I,J) / G(J)
   CALL AXBS (S, A, W, -NG, NF, NV, N1)
   CALL PRTS (W, NG, NF, 5HCENT., N1)
   CALL PRTS (C, NV, NF, 6HCOREL., N1)
   DFW = TN - GN
   PRINT 105, GM, DFW
105 FORMAT (// 26H UNIVARIATE F-TESTS. DFB =, F3.0,
  16H DFW =, F6.0 /// 18H VARIABLE  F-RATIO, 6X, 1HP)
   DO 115 I = 1,NV
   B = 0.0
   DO 110 J = 1,NG
110 B = B + S(I,J)**2 * G(J)
   CC = T(I)**2 / TN
   F = ((B - CC) * DFW) / ((Q(I) - B) * GM)
   P = PRBF(GM, DFW, F)
115 PRINT 120, I, F, P
120 FORMAT (/ I6, F12.4, F10.4)
   CALL PRTS (S, NV, NG, 6HG MEAN, N1)
   IF (KT .EQ. 0) GO TO 5
   REWIND 2      $ NT = TN
   DO 130 I = 1,NT
   READ (2) ID, (X(J), J = 1,NV)
   DO 125 J = 1,NF
125 Y(J) = SCPF(X, A, 1, J, NV, N1)
130 CALL SUBS (Y, NF, 2HDS, ID, NDS)
   GO TO 5   $  RETURN
   END
```

## Subroutine DMG

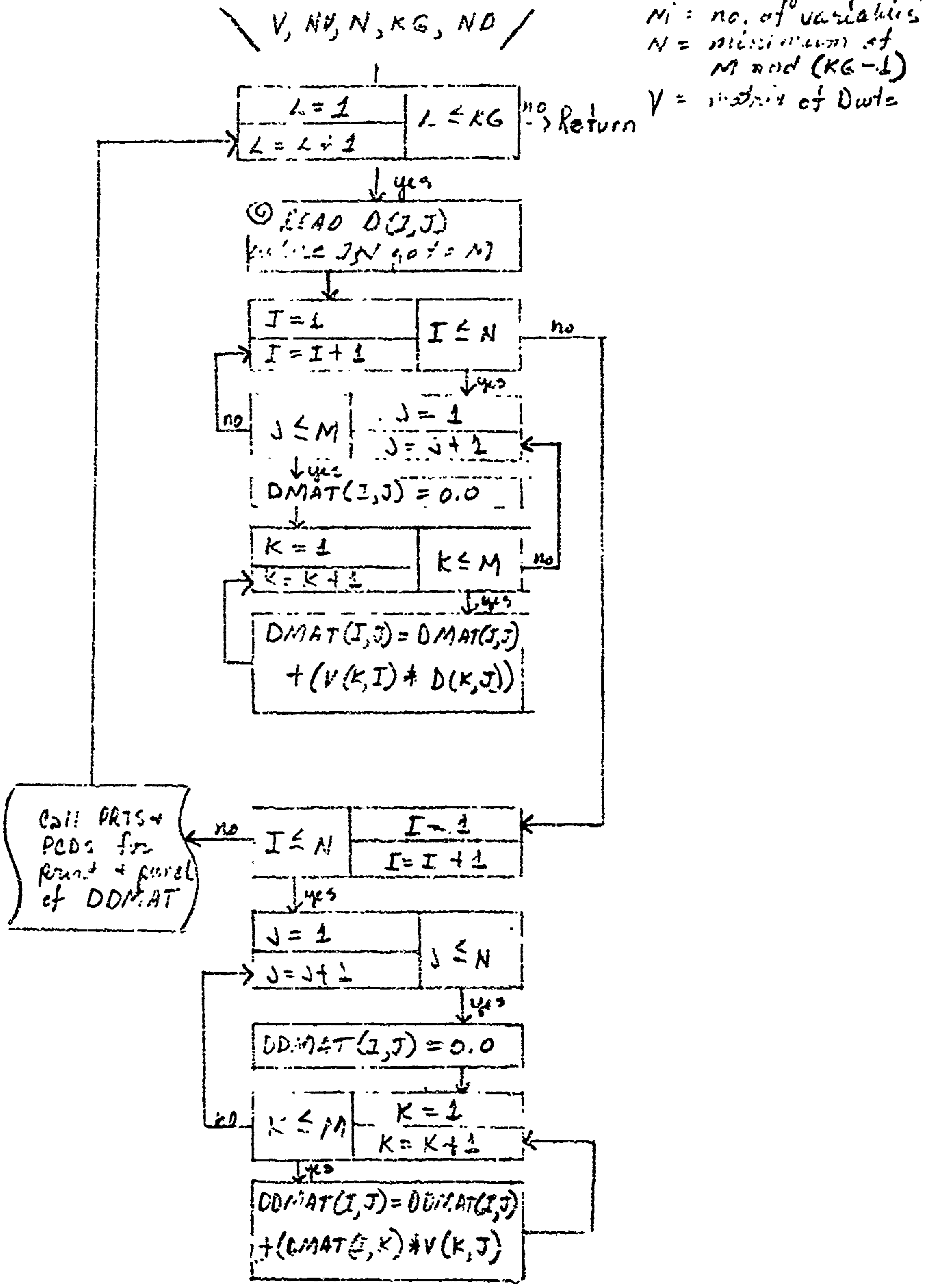
This 'dispersion matrix generator' subroutine is based on the RSPACE program of Cooley and Lohnes which requires the input of the discriminant function weights, vectors of group means and group dispersion matrices, and outputs centroids of the groups and dispersion of groups in reduced space. As our DISCRIM program produces centroids, the subroutine is necessary for the calculation of the reduced space dispersion matrices only.

The DWTS are input to the subroutine in the call using the variable A, the original space matrices are read within the subroutine from the rewound s  tch tape 1.

"The r x r dispersion matrix $DD_g$ in the discriminant space for group g may be obtained by pre- and post-multiplying the test space dispersion matrix $D_g$ for the group g by the matrix V containing the discriminant function (weights) vectors as follows:

$$DD_{g(r,r)} = V^1{}_{(r,m)} \cdot D_{g(m,m)} \cdot V_{(m,r)}."$$

Subroutine DMG



```
      SUBROUTINE DMG(V, NV, N, KG, ND)
      DIMENSION V(ND,N),  D(70,70),  DMAT(70,70),  DDMAT(70,70)
      DIMENSION V(ND,N), D(30,30), DMAT(30,30),DDMAT(30,30)
      PRINT 50,NV,N,KG,ND
   50 FORMAT (10X,6I10)
      REWIND 1
      DO 2 I=1,NV
    2 PRINT 7,(V(I,J),J=1,N)
      DO 23 L=1,KG
C     PRINT 18
   18 FORMAT (10X,* DISPERSION MATRIX OF INITIAL DATA*)
      DO 19 I= 1,NV
   19 READ  (1) (D(I,J),J=1,NV)
    7 FORMAT (10X,10F10.4)
      DO 20 I=1,N
      DO 20 J=1,NV
      DMAT(I,J)=0.0
      DO 20 K=1,NV
   20 DMAT(I,J) = DMAT(I,J)  + (V(K,I)*D(K,J))
      DO 21 I=1,N
      DO 21 J=1,N
      DDMAT(I,J)=0.0
      DO 21 K=1,NV
   21 DDMAT(I,J) = DDMAT(I,J) +  (DMAT(I,K)*V(K,J))
      PRINT 22
      CALL PCDS(DDMAT,N,N,5H DM=2,ND)
      DO 60 I=1,N
   60 PRINT 7,(DDMAT(I,J),J=1,N)
   22 FORMAT(10X,*  DISPERAION MATRIX OF REDUCED SPACE*)
   23 CONTINUE
      REWIND 1
      RETURN   $   END
```

## CLASSID - TPCLASS

The classification routines described by Cooley and Lohmes in Chapter seven of their book "Multivariate Procedures for the Behavioral Sciences" from the basis for the classification of sleep stages in the programs CLASSID and TPCLASS. The CLASSID program is a direct modification of the CLASSIF routine. The modified version is compatible with the format of the training set data produced for the sleep study and produces a summary of the classification results for each group rather than the output of a record by record classification as does CLASSIF.

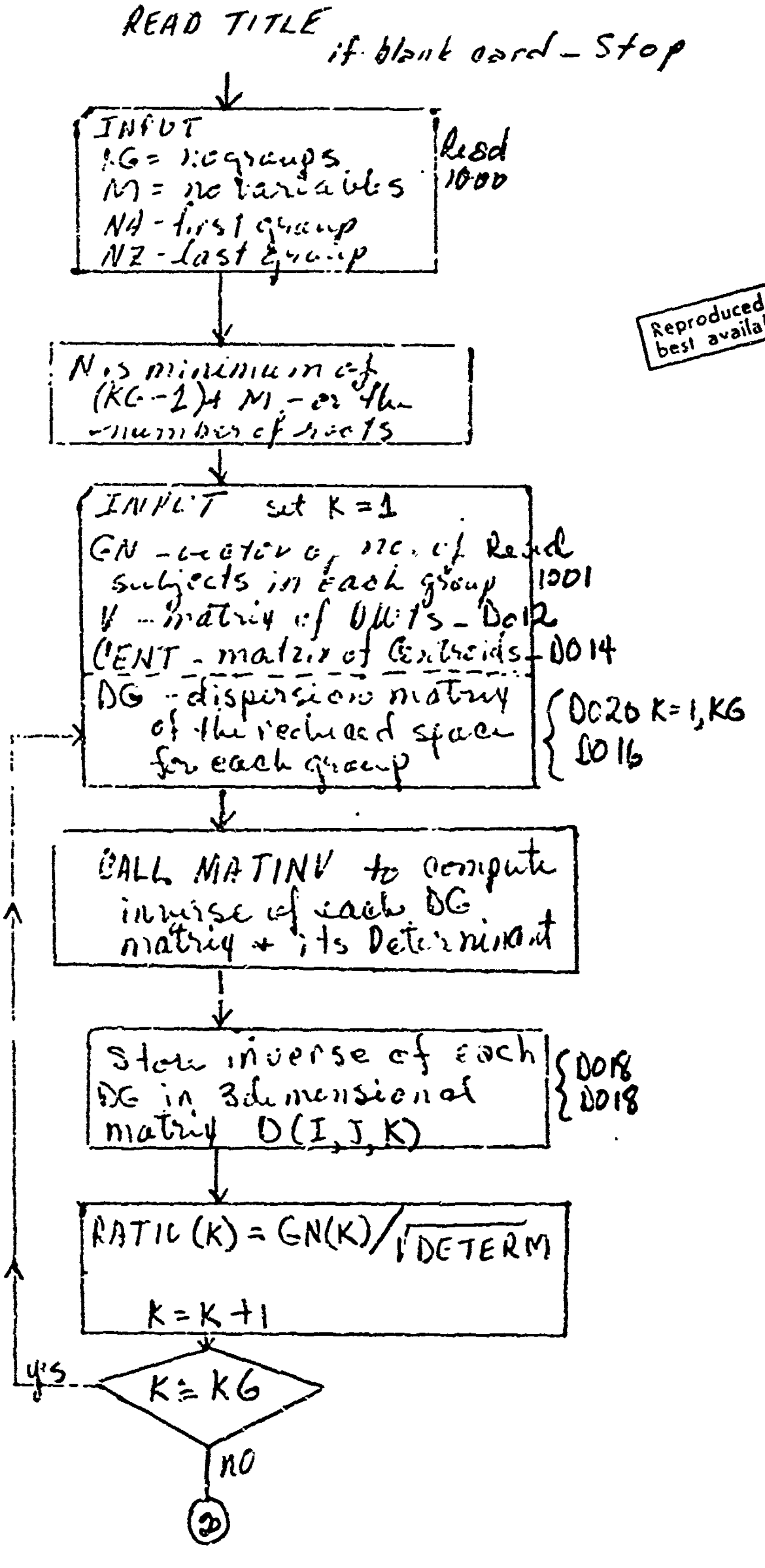
The program TPCLASS is a further extension of this probability classification system, taking as data input, a tape of unordered, non grouped sleep measures. The TPCLASS program produces a plotting of the probability classified scores versus the time of each data record. It also provides an accounting of the probability classified score as compared with the manually classified stage which is available on the data tape used.

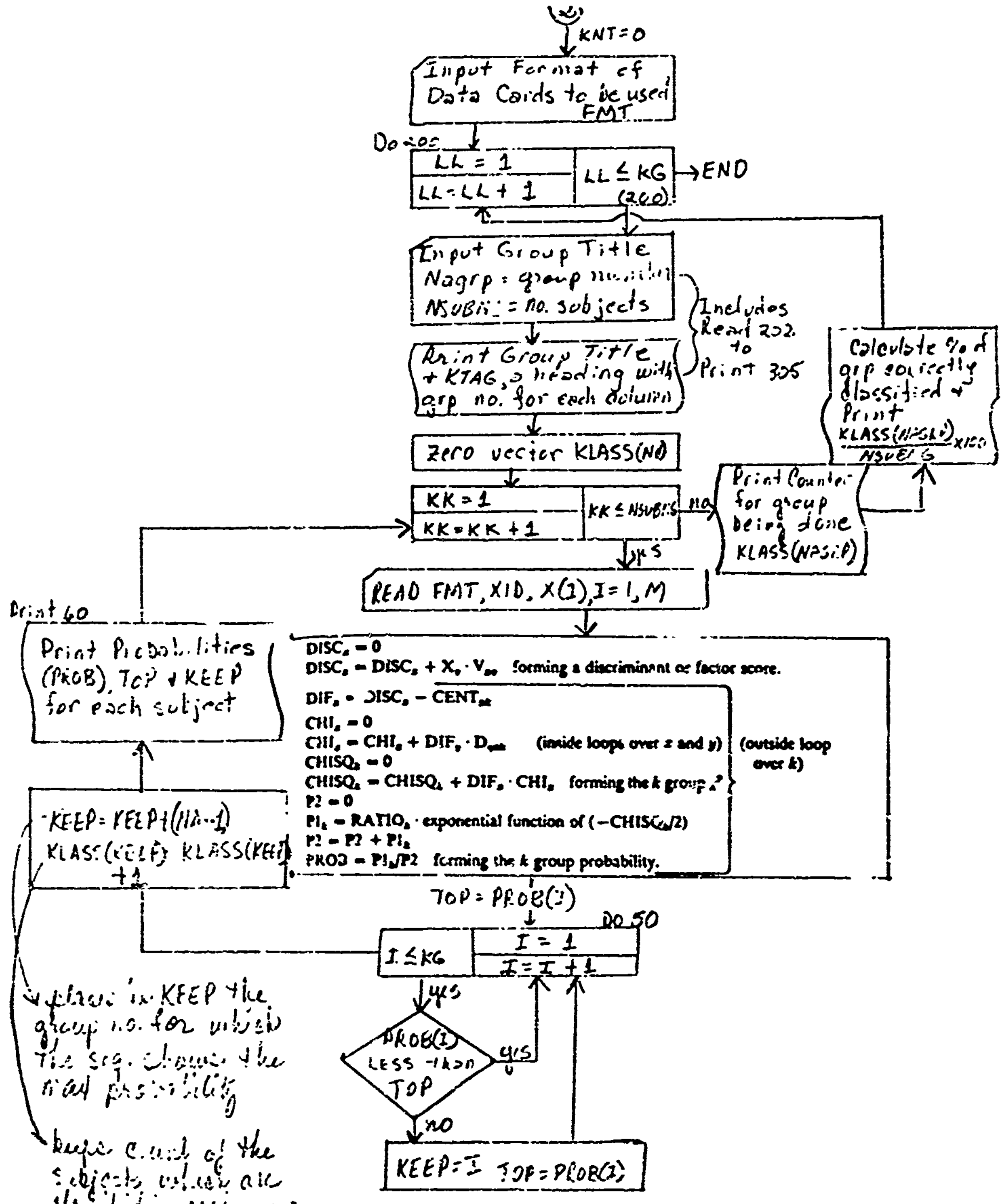
In both programs the subroutine MATINV is that of Cooley and Lohmes, a matrix inversion and determinant calculation by the Gauss-Jordan Method described in Chapter nine of the above mentioned work.

```
      PROGRAM CLASSID(INPUT,OUTPUT,TAPE1,TAPE2,PUNCH)
C     A=  TITLE  -- 12A6
C     B=  NO GROUPS,NO VARIABLES,FIRST GRP NO,LAST GRP NO -- 5I2
C     C=  NO SUBJECTS IN EACH GROUP  --  20F4.0
C     D=  D. WEIGHTS -- FORMAT 1005
C     E=  CENTROIDS
C     F=  DISPERSION MATRICES
C     G=  FORMAT OF DATA TO BE CLASSIFIES -- 12A6
C     H=  DATA WITH GROUP DIVIDER CARDS -- 2I5  (NO IN GRP , NO OF GRP)
      DIMENSION V(50,20),CENT(20,20), DG(50,50), D(20,20,20),RATIO(20),
     1GN(20), X(50), DISC(20), DIF(20), CHI(20),CHISQ(20), P1(20),
     2PROB(20),B(50),KLASS(20),KTAG(20),I6(6)
      INTEGER FMT(24) ,XID(1)
      DATA(I6(I),I=1,6)/10H(1X,A5,I5,,10H1X,        ,10H          ,
     110H(F5.3,1X),,10H*(*F4.3*) ,10HIN *12)    /
      DIMENSION INAME(12)
1000  FORMAT (10I2)
1001  FORMAT (20F4.0)
1002  FORMAT (5E14.7)
1003  FORMAT (12A6)
1004  FORMAT (10X,E14.7)
1005  FORMAT (10X,7F10.4)
1010  FORMAT(1X,12A6)
5000  READ 1003,INAME
      IF(INAME(1).EQ.INAME(2)) STOP
      PRINT 1010,INAME
      KNT = 0
      REWIND 1
      REWIND 2
      READ 1000,KG,M,NA,NZ
      N = MINO(NZ-1,M)
      ENCODE (10,2,I6(3))KG
    2 FORMAT (I2,8H        )
      PRINT 3,KG
    3 FORMAT (10X,*KG=  *,I10)
C     VECTOR LISTING NO SUBJECTS IN EACH GROUP
      READ 1001,(GN(NN),NN=1,KG)
      PRINT 1015
 1015 FORMAT(5X,*   NUMBER OF SUBJECTS IN EACH GROUP*)
      PRINT 1005,(GN(NN),NN=1,KG)
C  MATRIX OF D WEIGHTS
      DO 12 I =1,M
   12 READ 1005, (V(I,J), J=1,N)
C  MATRIX OF CENTROIDS
      DO 14 K= 1,KG
   14 READ 1005, (CENT(I,K), I = 1,N)
C     READ 1005,(CENT(1,K),K=1,KG)
C     DISPERSION MATRIX
      DO 20 K= 1,KG
      DO 16 I = 1,N
   16 READ 1005, (DG(I,J), J=1,N )
      CALL MATINV(DG,N,B, 0, DETERM  )
      DO 18 I=1,N
      DO 18 J=1,N
   18 D(I,J,K) = DG(I,J)
   20 RATIO (K) = GN(K)/ SQRTF (DETERM)
      PRINT 495 ,(RATIO(K),K=1,KG)
  495 FORMAT (1X,*  RATIO(K)  *,10E12.4/10E12.4)
      READ 1003,(FMT(I),I=1,12    )
```

```
      DO 200 LL=1,KG
      READ 202, NSUBNG,NAGRP
  202 FORMAT (3I5)
      PRINT 203,NAGRP,NSUBNG
  203 FORMAT(1H1,*GROUP * I5 * NUMBER SUBJECTS =    *I5)
      DO 303 MM=NA,NZ
  303 KTAG(MM)= MM
      PRINT 304,(KTAG(MM),MM=NA,NZ)
  304 FORMAT(7X,*GROUP*,1X,16(I3,3X))
      PRINT 305
  305 FORMAT(1X,* ID   COUNT*)
      DO 201 I= 1,16
  201 KLASS(I) = 0
      DO 300 KK=1,NSUBNG
      READ  FMT ,XID,(X(I),  I=1,M)
      KNT = KNT + 1
      DO 24 J=1,N
      DISC(J) = 0.0
      DO 24 I=1,M
   24 DISC(J) = DISC(J)+  (X(I)*V(I,J))
      DO 31 K=1,KG
      DO 28 I=1,N
   28 DIF(I) = DISC(I)-CENT(I,K)
      DO 30 J=1,N
      CHI(J) = 0.0
      DO 30 I=1,N
   30 CHI(J) = CHI(J)+(DIF(I)*D(I,J,K))
      CHISQ(K) = 0.0
      DO 31 I=1,N
   31 CHISQ(K) = CHISQ(K) + (DIF(I)*CHI(I))
C     WRITE (1) KNT,XID,(CHISQ(I),I=1,KG)
   33 P2 = 0.0
      DO 34 K=1,KG
      P1(K) = RATIO(K) * EXPF(-CHISQ(K)/2.0)
   34 P2 = P2+P1(K)
      DO 36 K=1,KG
   36 PROB(K) = P1(K)/P2
      TOP = PROB(1)
      KEEP = 1
      DO 50 I= 1,KG
      IF(PROB(I).LT.TOP)GO TO 50
      KEEP = I
      TOP = PROB(I)
   50 CONTINUE
      KEEP = KEEP +(NA-1)
      KLASS( KEEP) = KLASS(KEEP) + 1
      PRINT 16,XID,KNT,(PROB(I),I=  KG),TOP,KEEP
  300 CONTINUE
      PRINT 800,KLASS(NAGRP)
  800 FORMAT (10X,*KLASS(NAGRP)   =   *, I5)
      PCTCOR = (KLASS(NAGRP)*100.0)/NSUBNG
      PRINT 70,PCTCOR
   70 FORMAT(10X,*PERCENT CORRECTLY CLASSIFIED = * F10.5/)
  200 CONTINUE
      GO TO 5000
      CALL EXIT  $ END
```

# CLASSID


READ TITLE
if blank card – Stop
INPUT
KG = no groups
M = no variables
NA - first group
NZ - last group
Read 1000
N is minimum of (KG-1) + M - or the number of roots
INPUT set K = 1
GN - vector of no. of subjects in each group
Read 1001
V - matrix of UMTs - D012
CENT - matrix of Centroids - D014
DG - dispersion matrix of the reduced space for each group
D020 K=1,KG
D016
CALL MATINV to compute inverse of each DG matrix + its Determinant
Store inverse of each DG in 3 dimensional matrix D(I,J,K)
D018
D018
RATIO (K) = GN(K)/√DETERM
K = K + 1
K = KG
yes
NO
2


KNT=0
Input Format of Data Cards to be used FMT
Do 200
LL = 1
LL=LL + 1
LL ≤ KG (200)
END
Input Group Title
Nagrp = group number
NSUBJS = no. subjects
Includes Read 202 to Print 305
Print Group Title + KTAG, a heading with grp no. for each column
Calculate % of grp correctly classified + Print
Zero vector KLASS(NG)
KK = 1
KK=KK + 1
KK ≤ NSUBJS
no
yes
Print Counter for group being done KLASS(NASIP)
READ FMT, XID, X(I), I= 1, M
Print 60
Print Probabilities (PROB), TOP + KEEP for each subject
DISC_s = 0
DISC_s = DISC_s + X_v · V_vs forming a discriminant or factor score.
DIF_s = DISC_s − CENT_sk
CHI_s = 0
CHI_s = CHI_s + DIF_y · D_syk (inside loops over z and y)
(outside loop over k)
CHISQ_k = 0
CHISQ_k = CHISQ_k + DIF_s · CHI_s forming the k group χ²
P2 = 0
PI_k = RATIO_k · exponential function of (−CHISQ_k/2)
P2 = P2 + PI_k
PROB = PI_k/P2 forming the k group probability.
KEEP = KEEP+(NA−1)
KLASS(KEEP) = KLASS(KEEP) + 1
TOP = PROB(1)
DO 50
I = 1
I = I + 1
I ≤ KG
yes
PROB(I) LESS than TOP
yes
no
KEEP = I TOP = PROB(I)
places in KEEP the group no. for which the sub. shows the max probability
keeps count of the subjects which are classified in each group

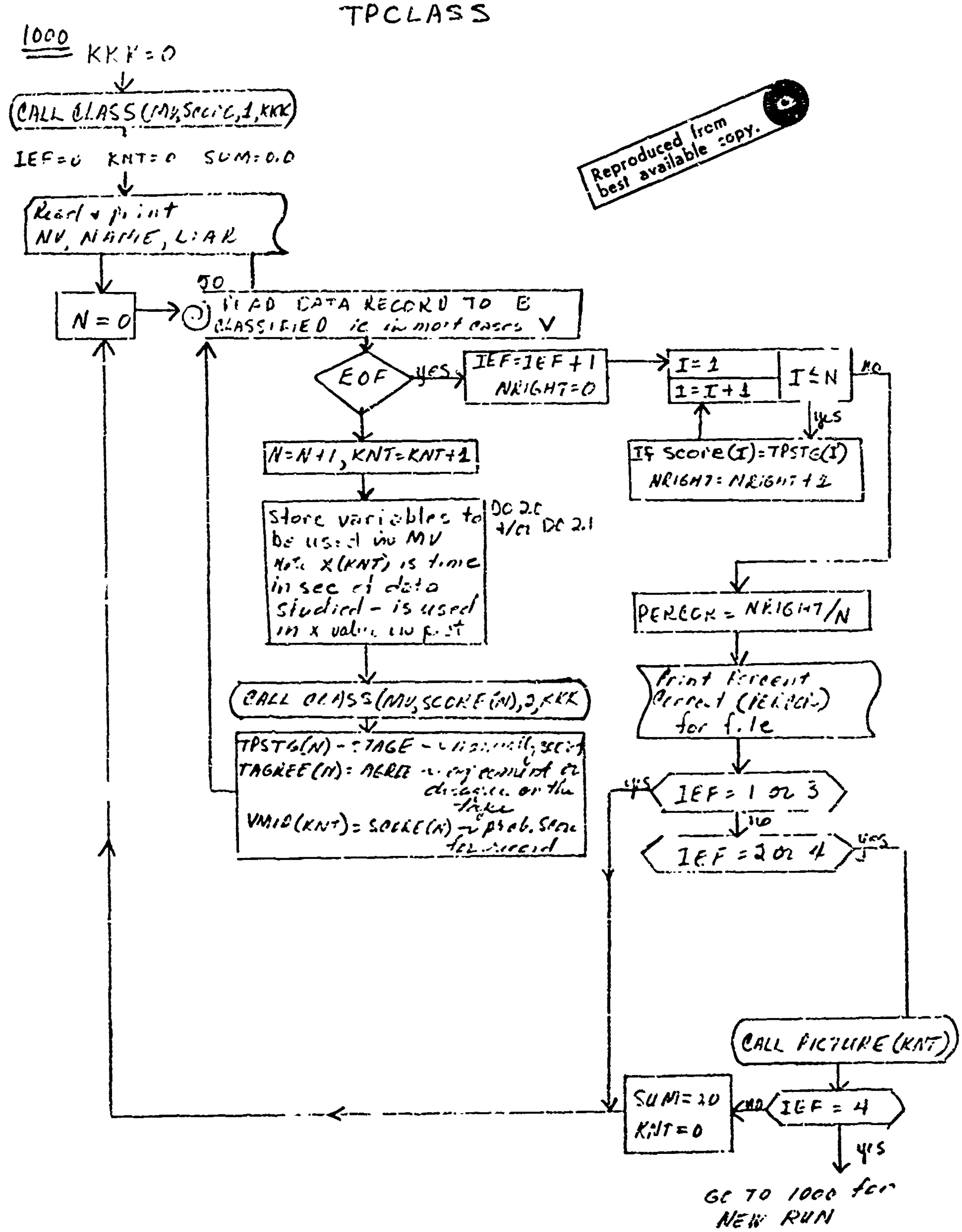
```
      SUBROUTINE MATINV(A,N,B,M,DETERM)
      DIMENSION IPIVOT(50),A(50,50),B(50,1),INDEX(50,2),PIVOT(50)
      COMMON PIVOT,INDEX,IPIVOT
      EQUIVALENCE(IROW,JROW),(ICOLUM,JCOLUM),(AMAX,T,SWAP)
      DETERM = 1.0
   15 DO 20 J=1,N
   20 IPIVOT(J)=0
   30 DO 550 I=1,N
C     SEARCH FOR PIVOT VALUE
   40 AMAX=0.0
   45 DO 105 J=1,N
   50 IF(IPIVOT(J)-1)60,105,60
   60 DO 100 K=1,N
   70 IF(IPIVOT(K)-1)80,100,740
   80 IF(ABSF(AMAX)-ABSF(A(J,K)))85,100,100
   85 IROW=J
   90 ICOLUM=K
   95 AMAX=A(J,K)
  100 CONTINUE
  105 CONTINUE
  110 IPIVOT(ICOLUM)=IPIVOT(ICOLUM)+1
C       INTERCHANGE ROWS TO PUT PIVOT ELEMENT ON DIAGONAL
  130 IF(IROW-ICOLUM)140,260,140
  140 DETERM=-DETERM
  150 DO 200 L=1,N
  160 SWAP=A(IROW,L)
  170 A(IROW,L)=A(ICOLUM,L)
  200 A(ICOLUM,L)=SWAP
  205 IF(M)260,260,210
  210 DO 250 L=1,M
  220 SWAP=B(IROW,L)
  230 B(IROW,L)=B(ICOLUM,L)
  250 B(ICOLUM,L)=SWAP
  260 INDEX(I,1)=IROW
  270 INDEX(I,2)=ICOLUM
  310 PIVOT(I)=A(ICOLUM,ICOLUM)
  320 DETERM=DETERM*PIVOT(I)
C         DIVIDE PIVOT ROW BY PIVOT ELEMENT
  330 A(ICOLUM,ICOLUM)=1.0
  340 DO 350 L=1,N
  350 A(ICOLUM,L)=A(ICOLUM,L)/PIVOT(I)
  355 IF(M)380,380,360
  360 DO 370 L=1,M
  370 B(ICOLUM,L)=B(ICOLUM,L)/PIVOT(I)
C       REDUCE NON-PIVOT ROWS
  380 DO 550 L1=1,N
  390 IF(L1-ICOLUM)400,550,400
  400 T=A(L1,ICOLUM)
  420 A(L1,ICOLUM)=0.0
  430 DO 450 L=1,N
  450 A(L1,L)=A(L1,L)-A(ICOLUM,L)*T
  455 IF(M)550,550,460
  460 DO 500 L=1,M
  500 B(L1,L)=B(L1,L)-B(ICOLUM,L)*T
  550 CONTINUE
C       INTERCHANGE COLUMNS
  600 DO 710 I=1,N
  610 L=N+1-I
  620 IF(INDEX(L,1)-INDEX(L,2))630,710,630
```

```
630 JROW=INDEX(L,1)
640 JCOLUM=INDEX(L,2)
650 DO 705 K=1,N
660 SWAP=A(K,JROW)
670 A(K,JROW)=A(K,JCOLUM)
700 A(K,JCOLUM)=SWAP
705 CONTINUE
710 CONTINUE
740 RETURN
    END
```

```
      PROGRAM TPCLS(INPUT,OUTPUT,TAPE1,PUNCH)
CCCCCC(
CCCCCC
C DATA CARDS.    A=KG,M (3I2)
C                B= GN(KG)  ..(20F4.0)
C                C= DWTS(V(I,J))    (10X,6F10.4)
C                D=  CENTROIDS (CENT(I,K))    (10X,6F10.4)
C                E=D MAT(DG(I,J)     (10X,7F10.4)
C                F= NV  ALSO TITLE    (I5,6A10)
C                G= LIST OF VARIABLES (2X,20I3)
CCCCCC
CCCCC
      COMMON/PP/VMID(300),NAME(4),TPS(300),X(300)
      DIMENSION TIMES(6),V(22),SCORE(200),TPSTG(200),TAGREE(200),
     1LVAR(22) ,XX(150),Z(11)
      REAL MV(33)
      ASSIGN 58 TO NNN
      CALL XMIT(NNN)
      CALL BGNPLT(4LPLOT  ,   10.0,50,50)
 1000 KKK = 0
      PRINT 1
    1 FORMAT (1H1)
      CALL CLASS(MV,SCORE(1),1,KKK)
    9 CONTINUE
      IEF = 0  $ KNT = 0
      SUM = 0.0
C    NUMBER VARIABLES TO BE USED (I5) --ANY TITLE YOU WISH
      READ 2,NV,(NAME(I),I=1,6)
    2 FORMAT (1X,I4,6A10)
      PRINT 2,NV,(NAME(I),I=1,6)
    6 FORMAT(10X,*SEQUENCE OF VARIABLES TO  BE USED*)
      READ  5,(LVAR(I),I=1,NV)
    5 FORMAT (2X,20I3)
      PRINT 6
      PRINT 5,(LVAR(I),I=1,NV)
   70 N = 0
      PUNCH 300,NAME
  300 FORMAT(4A10)
      PRINT 7
    7 FORMAT(1H1,1X,*FILE BEGINS WITH THIS RECORD*)
   50 CONTINUE
      READ (1) NT,TIMES,STAGE,AGREE,V  ,DOMEGA,TAU,Z
      IF(ENDFILE 1) 30,11
   11 IF(N.EQ.0) PRINT 12 ,        STAGE,AGREE,V ,Z
   12 FORMAT(1X,                              F5.0,A15,2X,*PERIODS*,
     12F10.3,9F3.0/ 76X,*RATES*,2F10.3,9F3.0,/1X,11F10.4)
      N = N + 1
      KNT = KNT + 1
      DO 20 L=1,NV
      KEEP =LVAR(L)
   20 MV(L) = V(KEEP)
      DO 21 L = 1,11
      KK = L + 9
   21 MV(KK) = Z(L)
      X(KNT) = (V(1)/1000.0)*128.0+ SUM
      XX(N) = X(KNT)
      SUM = X(KNT)
      SCORE(N)=0.0
      CALL CLASS(MV,SCORE(N),2,KKK)
```

```
      TPSTG (N) = STAGE  $ TAGREE(N) = AGREE
      TPS(KNT) = STAGE
      VMID(KNT) = SCORE(N)
      GO TO 50
   30 IEF = IEF+1
      NRIGHT = 0
      DO 31 J = 1,N
      IF(SCORE(J).EQ.TPSTG(J)) NRIGHT = NRIGHT+1
      NCHK = 10H
      IF(SCORE(J).NE.TPSTG(J)) NCHK = 10H MISSED
   31 CONTINUE
   32 FORMAT(1X,I3,2F10.0,2(5X,A10),5X,E13.3)
      RIGHT = NRIGHT  $  COUNT = N
      PERCOR = (RIGHT/COUNT)*100.0
      PRINT 34,IEF,PERCOR
   34 FORMAT(10X,*FILE  *,I2,*  PERCENT CORRECT =  *,F5.2)
      PUNCH 301,IEF,PERCOR
  301 FORMAT(1X,*FILE*,I5,F10.3)
      GO TO(70,80,70,80),IEF
   80 CONTINUE
      CALL PICTURE(KNT)
      IF(IEF.EQ.4) GO TO 60
      SUM = 0.0
      KNT = 0  $  GO TO 70
   58 PRINT 57,IEF,N
   57 FORMAT(1X,*XMIT IN CONTROL,FILE*,I5,*  RECORD*,I5)
      PRINT 12,STAGE,AGREE,V,Z
      STOP
   60 GO TO 1000
      END
```

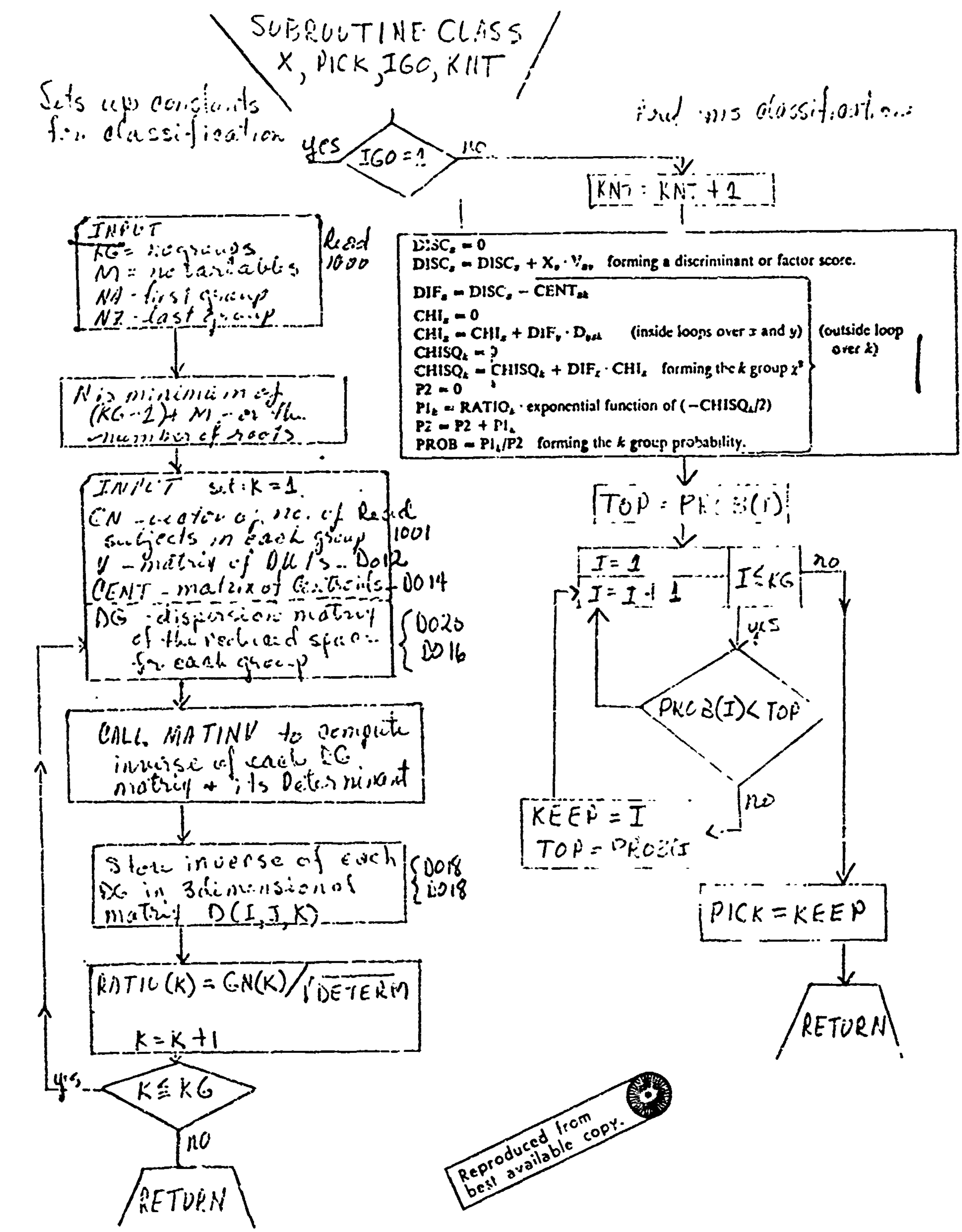
TPCLASS



1000 KKK=0
CALL CLASS(MV,Score,1,KKK)
IEF=0 KNT=0 SUM=0.0
Read + print NV, NAME, LIAR
N=0
50
READ DATA RECORD TO BE CLASSIFIED ie in most cases V
EOF
yes
IEF=IEF+1 NRIGHT=0
I=1
I=I+1
I≤N
no
yes
IF Score(I)=TPSTG(I) NRIGHT=NRIGHT+1
N=N+1, KNT=KNT+1
Store variables to be used in MV
Note X(KNT) is time in sec of data studied - is used in x value in plot
DC 2.0
or DC 2.1
PERCOR = NRIGHT/N
Print Percent Correct (PERCOR) for file
CALL CLASS(MV,SCORE(N),2,KKK)
TPSTG(N) - STAGE
TAGREE(N): AGREE
VMID(KNT) = SCORE(N)
IEF = 1 or 3
yes
no
IEF = 2 or 4
yes
CALL PICTURE(KNT)
SUM=0.0 KNT=0
no
IEF = 4
yes
GO TO 1000 for NEW RUN

```
      SUBROUTINE CLASS (X,PICK,IGO,KNT)
      DIMENSION V(50,20),CENT(20,20), DG(50,50), D(20,20,20),RATIO(20),
     1GN(20), X(50), DISC(20), DIF(20), CHI(20),3(50),KLASS(20),KTAG(20)
     2,CHISQ(20)    ,PROB(20),P1(20)
      INTEGER FMT(24)
 1000 FORMAT (3I2)
 1001 FORMAT (20F4.0)
 1002 FORMAT (5E14.7)
 1003 FORMAT (12A6)
 1004 FORMAT (10X,E14.7)
 1005 FORMAT (10X,7F10.4)
      IF (IGO.EQ.1) GO TO 100
   10 KNT = KNT+ 1
      DO 24 J=1,N
      DISC(J) = 0.0
      DO 24 I=1,M
   24 DISC(J) = DISC(J)+  (X(I)*V(I,J))
      DO 31 K=1,KG
      DO 28 I=1,N
   28 DIF(I) = DISC(I)-CENT(I,K)
      DO 30 J=1,N
      CHI(J) = 0.0
      DO 30 I=1,N
   30 CHI(J) = CHI(J)+(DIF(I)*D(I,J,K))
      CHISQ(K) = 0.0
      DO 31 I=1,N
   31 CHISQ(K) = CHISQ(K) + (DIF(I)*CHI(I))
   33 P2 = 0.0
      DO 34 K=1,KG
      ZAVE = EXPF(-CHISQ(K)/2.0)
      P1(K) = RATIO (K) * ZAVE
   34 P2 = P2 +P1(K)
      DO 36 K=1,KG
   36 PROB(K) = P1(K)/P2
      TOP = PROB(1)
      KEEP = 1
      DO 50 I= 1,KG
      IF(PROB(I).LT.TOP)GO TO 50
      KEEP =I $ TOP = PROB(I)
   50 CONTINUE
      PICK =  KEEP - 1
      RETURN
   60 PICK = 0.0
      PRINT 61,KNT,P2
   61 FORMAT (10X,*EPOCH NO.  *,I4,*  P2 =  *,D13.5)
      RETURN
  100 KNT = 0
      READ 1000, KG, M  ,N1
      IF(.NOT.KG)101,80
  101 N = MINO(KG-1,M)
      PRINT 102,KG,M,N1
  102 FORMAT(1X,3I2)
      REWIND 1
C     VECTOR LISTING NO SUBJECTS IN EACH GROUP
      READ  1001,(GN(NN),NN=1,KG)
      PRINT 1005,(GN(NN),NN=1,KG)
C  MATRIX OF D WEIGHTS
      DO 12 I =1,M
   12 READ 1005,(V(I,J),  J=1,N)
```

```
C  MATRIX OF CENTROIDS
      DO 14 K= 1,KG
   14 READ 1005, (CENT(I,K), I = 1,N)
C    DISPERSION MATRIX
      DO 20 K= 1,KG
      DO 16 I = 1,N
   16 READ  1005,  (DG(I,J),  J=1,N)
      CALL MATINV(DG,N,B, 0, DETERM  )
      PRINT 301,DETERM
  301 FORMAT (20X,*DETERM=   *,E13.5)
   17 CONTINUE
      DO 18 I=1,N
      DO 18 J=1,N
   18 D(I,J,K) = DG(I,J)
   20 RATIO (K) = GN(K)/ SQRTF (DETERM)
      PRINT 495 ,(RATIO(K),K=1,KG)
  495 FORMAT (1X,*  RATIO(K)  *,10E12.4/10E12.4)
      RETURN
   80 CALL ENDPLT  $  STOP
      END
```


SUBROUTINE CLASS
X, PICK, IGO, KNT
Sets up constants for classification
Find misclassification
yes
IGO = 1
no
KNT = KNT + 1
INPUT
KG = no groups
M = no variables
N1 - first group
N2 - last group
Read 1000
N is minimum of (KG-1) + M - or the number of roots
INPUT set K = 1
CN - vector of no. of subjects in each group
Read 1001
V - matrix of DU's
D012
CENT - matrix of Centroids
D014
DG - dispersion matrix of the reduced space for each group
D020
D016
CALL MATINV to compute inverse of each DG matrix + its Determinant
Store inverse of each DG in 3 dimensional matrix D(I,J,K)
D018
D018
RATIO(K) = CN(K)/√DETERM
K = K + 1
yes
K ≤ KG
no
RETURN
DISC_s = 0
DISC_s = DISC_s + X_v · V_sv forming a discriminant or factor score.
DIF_s = DISC_s − CENT_sk
CHI_s = 0
CHI_s = CHI_s + DIF_y · D_ysk (inside loops over x and y)
(outside loop over k)
CHISQ_k = 0
CHISQ_k = CHISQ_k + DIF_x · CHI_x forming the k group χ²
P2 = 0
P1_k = RATIO_k · exponential function of (−CHISQ_k/2)
P2 = P2 + P1_k
PROB = P1_k/P2 forming the k group probability.
TOP = PROB(I)
I = 1
I = I + 1
I ≤ KG
no
yes
PROB(I) < TOP
no
KEEP = I
TOP = PROB(I)
PICK = KEEP
RETURN

## Subroutine Picture

Loops 1 & 2 set up a matrix* showing how the computer score of each record compares with the manual scores read from the tape. VMID is computer score, TPS is manual score.

Loop 8 punches (and prints) this matrix of comparison for use in correlation program Correls.

Loop 10 converts the computer score to an appropriate value for plotting i.e. stage 0 becomes 5, 1 is 4, 2 is 3, 3 is 2, 4 is 1, and 5 is 0. Stage 5 or REM scores are converted to a dotted line between stages 0 and 1.

The statements from 10 on perform the plotting according to the plot routines available for the CDC 6600 computer at UT at Austin.

* This is a very compact way of saving two scores for each record of data processed in X(I,J) where X is count of the no of records manually classified as I, which were computer classified as J.

```
      SUBROUTINE PICTURE (NP)
      COMMON/PP/VMID(300),NAME(4),TPS(300),X(300)
      DIMENSION Y(300),IPT(300),ISET(6,6)
      DO 1 L = 1,6
      DO 1 J = 1,6
    1 ISET(L,J) = 0
    4 DO 2 L = 1,NP
      J = TPS(L) + 1   $ K = VMID(L) + 1
      IF(J.EQ.10) GO TO 2
      ISET(J,K) = ISET(J,K) + 1
    2 CONTINUE
      PRINT 11
      PRINT 76
   76 FORMAT (15X,*MANUALLY SCORED STAGE*)
      PRINT 77
   77 FORMAT(10X,*  0    1    2    3    4    5    *)
      DO 8 J =1,6
      JJ = J -1
      PUNCH 99,(ISET(K,J),K=1,6)
   99 FORMAT(6I3)
    8 PRINT 9,JJ,(ISET(K,J),K = 1,6)
    9 FORMAT(4X,I1,5X,6(I4,1X))
      PRINT 11
   11 FORMAT(1X,//)
      DO 5 K = 1,NP
      Y(K) = 4.2  $  IPT(K) =  3
    5 CONTINUE
      DO 10 J = 1,NP
      VMID(J) = ABS(VMID(J)-5)
      IF(VMID(J).NE.0.0) GO TO 10
      VMID(J) = 4.0  $  IPT(J) = 2
   10 CONTINUE
      SP = X(NP)/3600.0
      ENCODE(10,7,IXP,X)SP
    7 FORMAT(F5.2,5H HRS.  )
      ENCODE (10,17,INP)NP
   17 FORMAT(I5,5H PTS.   )
      CALL PLT(2.0,2.0,-3)
      X(NP+1) = 0.0  $  X(NP+2) = 5000.0
      VMID(NP+1) = 0.0  $ VMID(NP+2) = 1.6
      CALL LINE(X,VMID,NP,1,0,0)
      DO 90 J= 1,NP
      X(J) = X(J)/  X(NP + 2)
   90 CALL PLT(X(J),2.62,IPT(J))
      PLACE = 0.0  $  ICODE = -1
      DO 80 J = 1,7
      CALL SYMBOL(0.0,PLACE,0.14,9,0.0,ICODE)
      PLACE = PLACE + 0.625
      ICODE = ICODE -1
   80 CONTINUE
      PLACE = 0.0  $  ICODE = -1
      DO 81 L = 1,35
      CALL SYMBOL(PLACE,0.0,0.14,9,0.0,ICODE)
      PLACE = PLACE - 0.18
      ICODE = ICODE - 1
   81 CONTINUE
      PLACE = .625 $  INTEQ = 47
      DO 91 K = 1,5
      CALL SYMBOL(-1.4,PLACE,0.105,INTEQ,0.0,-1)
```

```
      PLACE = PLACE +.625 $ INTEQ = INTEQ - 1
   91 CONTINUE
      CALL SYMBOL(-1.6,1.55,0.14,11HSLEEP STAGE,90.0,11)
      CALL SYMBOL(2.0,-1.0,0.14,IXMAX,0.0,10)
      CALL SYMBOL(4.0,-1.0,0.14,INP  ,0.0,10)
      CALL SYMBOL(0.5,-1.50,0.14,NAME,0.0,40)
      CALL PLT(0.0,0.0,999)
      PRINT 20,NP,VMID(NP)
   20 FORMAT (10X,*PLOT COMPLETED--NO.POINTS=*,I8,* LAST PT.= *,F10.4)
      RETURN  $ END
```

Program Correls

DO 60 - sto. in matrix CODE(4,4) the cost values listed in XCODE(16) of the data statement. These values indicate the cost of error of the computer classification versus the manual classification.

100 thru DO 3 read in a classification matrix for one half night of data as prepared by program Tpclass.

DO 6 DO 7 DO 8 reduce the 6 x 6 Tpclass MATRIX to a 4 x 4 matrix by adding stage 1 to stage SREM and stage 3 to stage 4.

DO 50 Evaluate the e re score matrix by multiplying each value of the score matrix by the corresponding value of the cost matrix. Accumulate these products in the variable SUM. If every decision between sleep and awake were incorrect this sum would be equal to the number of data records classified.

DO 10 - 20 and 30 resort the 4 x 4 computer vs manual score matrix into two vectors of KNT length where KNT is no. of data records classified - store vectors on scratch tape 1 to be read back for intercorrelation analysis.

CALL CORSS (KNT, NV) to perform the Intercorrelation analysis. This routine is a modification of Veldmans' CORS subroutine to the extent that the data is input through the use of the scratch tape 1.

Print 50 - Print 53 - print sum as total error cost. Calculate and print error cost per data record classified.

DO 81 - Print 82 - The total of the records on the diagonal of the 4 x 4 score matrix represent correct classifcation. The percentage of these correct records versus the total number classified is calculated and output.

```
      PROGRAM CORRELS(INPUT,OUTPUT,TAPE1)
      DIMENSION XM(6,6),ITITLE(4) ,CODE(4,4)
      DIMENSION XCODE(16)
      DIMENSION NAME(8)
      DATA(XCODE(I),I=1,16)/0.,1.,1.,1.,1.,0.,.25,.5,1.,.25,0.,.25,
     11.,.50,.25,0.0/
      K = 0
      DO 60 I = 1,4
      DO 60 J = 1,4
      K = K + 1
      CODE(I,J) = XCODE(K)
   60 CONTINUE
      DO 61 K = 1,4
   61 PRINT 52,(CODE(K,L),L=1,4)
   52 FORMAT(1X,4F5.2)
      NV = 2
100   READ 1,IFILE1,PC1,ITITLE,IFILE2,PC2
    1 FORMAT(1X,/5X,I5,F10.0/4A10/5X,I5,F10.0)
      IF(ITITLE(1).EQ.ITITLE(2))CALL EXIT
      PRINT 2,ITITLE,IFILE1,IFILE2,PC1,PC2
    2 FORMAT(1H1,1X,4A10,*   FILES*,2I5,2F10.3)
      REWIND 1
      DO 3 I=1,6
      READ 4,(XM(I,J),J=1,6)
    4 FORMAT(6F3.0)
    3 PRINT 5,(XM(I,J),J=1,6)
    5 FORMAT(1X,6F5.0)
      DO 6 I = 1,6
      XM(I,2) = XM(I,2)+XM(I,6)
    6 XM(I,4) = XM(I,4)+XM(I,5)
      DO 7 J = 1,4
      XM(2,J) = XM(2,J) + XM(6,J)
    7 XM(4,J) = XM(4,J) + XM(5,J)
      PRINT 70
   70 FORMAT(1X,//,1X,*REDUCED MATRIX*)
      DO 8 K = 1,4
    8 PRINT 5,(XM(K,L),L=1,4)
      SUM = 0.0
      DO 50 I= 1,4
      DO 50 J= 1,4
   50 SUM = SUM + (XM(I,J)*CODE(I,J))
      KNT = 0
      DO 10 I = 1,4
      CS = I
      DO 20 J = 1,4
      SS = J
      L = XM(I,J)
      IF(L.EQ.0)GO TO 20
      DO 30 K = 1,L
      KNT = KNT + 1
      WRITE (1) KNT,CS,SS
C     PRINT 31,KNT,CS,SS
   31 FORMAT(1X,I5,2F5.0)
   30 CONTINUE
   20 CONTINUE
   10 CONTINUE
      REWIND 1
      CALL CORSS(KNT,NV)
      PRINT 51,SUM ,KNT
```

```
51 FORMAT(10X,//,10X,*ERROR COST =  *F10.2,* FOR * I5,* DATA POINTS*)
   COUNT = KNT
   CP = SUM / COUNT
   PRINT 53,CP
53 FORMAT(10X,*OR    *,F10.6,*  PER POINT*)
   DSUM = 0.0
   DO 81 I = 1,4
81 DSUM = DSUM + XM(I,I)
   PRINT 80,DSUM
80 FORMAT(10X,//,10X,F6.2,* SUM OF POINTS ON DIAGONAL*)
   PDSUM = (DSUM/COUNT)*100.0
   PRINT 82,PDSUM
82 FORMAT(10X,* OR  *,F10.4,*  PERCENT ON THE DIAGONAL*//)
   GO TO 100
   END
```